

IMPROVING CONSISTENCY IN TEACHER JUDGEMENTS

An Investigation for the Department of Education, Victoria

Marion Meiers, Clare Ozolins and Phillip McKenzie

FINAL REPORT

June 2007

Contents

Tables	iii
Acknowledgements	iv
Executive Summary	v
1. Introduction	1
2. Methodology	4
3. Evidence and Research Base	7
4. Current Practices in Australian States and Territories	17
5. International Perspectives	44
6. Findings	53
7. Conclusions and Recommendations	65
References	71
Additional Research and Policy Documents	73

TABLES

Table 1: Approaches to consistency of teacher judgements against state-wide standards in Australian states and territories	18
Table 2: Timeline for QCAR development	32
Table 3: Standards frameworks and key resources for developing consistency of teacher judgement in three international jurisdictions.....	45
Table 4: Key elements in a comprehensive strategy to improve consistency of teacher judgements	55
Table 5: Types of resources and illustrative costs in implementation of a comprehensive strategy in Victoria	63

ACKNOWLEDGEMENTS

The ACER Team

Ms Marion Meiers, Senior Research Fellow (Project Director)
Dr Phillip McKenzie, Acting Research Director, Teaching and Leadership Research Program
Ms Clare Ozolins, Research Officer
Ms Christine Rosicka, Research Officer
Ms Emma Curtin, Administrative Officer
The Cunningham Library

The ACER team wishes to acknowledge the support and assistance offered to us by curriculum teams and officers in all States and Territories. Special thanks to those in Tasmania, South Australia, and New South Wales who generously agreed to meet with us and shared valuable time, knowledge and resources. Dr Gabrielle Matters, from the ACER office in Brisbane, provided valuable advice.

We also acknowledge the support of the Department of Education team, Ms Madeleine Jenkins, Mr Maurice Sheehan and Mr Dionysios Koutsouvelis, who provided valuable assistance in collection of the Victorian data and advice on drafts.

EXECUTIVE SUMMARY

This project investigates current best-practice approaches to ensuring consistency of teacher judgements against P-10 state-wide standards. The investigation encompassed national and international approaches that had been implemented or were under development. The key purpose of the project was to identify and outline approaches that, on the basis of available evidence, are most effective in improving the consistency of teacher judgements against state-wide P-10 standards.

The issue of developing approaches to improve consistency of teacher judgements against state-wide standards is currently under discussion in all states and territories. The significance of teachers' judgements of their students' achievement is widely recognised. It is also recognised that the quality of these judgements is important, and that the consistency of judgements against a common standards framework might vary from school to school, unless there are common, shared understandings of the standards, and of the kind of evidence that indicates achievement against the standards. Thus there is an imperative for developing effective approaches to improving the consistency of teachers' judgements within schools, between schools, and over time.

In Victoria, a coherent approach linking to and building upon the Victorian Essential Learning Standards (VELS), progression points, continua, assessment maps and other policy, advice and support materials has been developed and implemented. This report suggests a number of ways to build on this base.

The overview of the research base was structured around four themes: assessment in education; standards-referenced assessment; moderation of teacher assessments; and teacher professional learning as a key strategy for improving the consistency of teacher judgements.

Five key elements of effective assessment systems identified through the research were used to review current Australian and overseas approaches:

- teachers' knowledge and understanding of the standards
- curriculum planning; opportunities for students to demonstrate achievement of standards
- evidence of student achievement
- assessment of the evidence against the standards
- validation of teachers' judgements.

In all Australian states and territories a range of strategies designed to improve the consistency of teachers' judgements of students' progress against common standards are in place. Some are large scale system-wide initiatives, and others are more localised. The most effective strategies combine state-wide and local initiatives into a strong framework. The development of these strategies is related to the development and continuing implementation of standards-based curriculum frameworks.

The investigation found strong links between best-practice approaches to building and improving the consistency of teacher judgements against state-wide standards and a strong focus on teacher professional learning.

A number of approaches were identified as being worthy of further consideration. They provide some signposts to the further development of effective approaches for improving the consistency of teacher judgement.

The development of consistency in teacher judgements involves a strengthening of classroom assessment practices in schools, in particular formative assessment practices. The investigation found that a shift towards increased emphasis on classroom assessment was prevalent in most jurisdictions. This was most effective where it was validated against external standards, whether by reference to state-wide data or other sources.

The provision of print and electronic resources such as annotated work samples, advice on moderation activities, assessment advice, and support documentation for state-wide standards play a key role in improving the consistency of teacher judgements, but on their own they appear to have only a limited impact.

Evidence on the extent of teacher consistency is difficult to obtain, and does not appear to be a focus of current efforts in Australia or overseas. The research evidence on the benefits and costs of effective approaches needs to be strengthened. The continuation of the current practice of comparing judgements of teachers' judgement against the VELs with AIM data will provide useful information to school leaders, and signal when there may be a need to refine teachers' judgements against the standards. In addition, some longitudinal studies of the judgements made by all teachers in selected schools, or all teachers at the same year level in several schools in different metropolitan and non-metropolitan locations, would help to generate further evidence of improvements in the consistency of teacher judgements. A longitudinal approach is the only appropriate way of gathering evidence of the consistency of judgements over time.

The investigation highlighted the significance of moderation processes in developing consistency of teacher judgement. It also identified the quality assurance function of moderation of teachers' judgements. The literature indicates that teachers' active participation in moderation activities within schools and across schools is a key element in developing consistency of teacher judgements. Moderation is recognised as an effective means of improving the quality of assessment in schools and of teachers' judgements, and offers significant benefits to the system as a whole.

Moderation processes involving teachers' collaborative examination of samples of student work should become a central aspect of schools' assessment culture and practices. Opportunities for teachers to engage in moderation activities within and between schools, in face-to-face and online forms should be implemented as a system-wide initiative. This needs to be complemented by an external means of monitoring teachers' assessments over time.

The Australian and international experience indicates the need to take a comprehensive approach to improving the consistency of teachers' judgements. No single element, no matter how well designed, will be sufficient on its own. The following recommendations are framed with a view to building up a high-quality framework in Victoria.

Recommendation 1

It is recommended that the key features of best-practice approaches to improving consistency of teacher judgements against statewide standards identified in the investigation be adopted in the Victorian context:

- a. processes for developing teachers' understanding of the standards*
- b. processes for assisting teachers to identify appropriate evidence of students' achievement of the standards*

- c. processes and protocols for teachers to meet together to examine samples of their own students' work against samples demonstrating achievement at the VELS standards and progression points, and to moderate their assessments of their own students' work*
- d. opportunities for moderation within schools and with other schools*
- e. leadership and support from trained assessment leaders available to all schools*
- f. processes for external validation of the accuracy of teachers' judgements after moderation*

Recommendation 2

It is recommended that the Department, in the light of the findings of the investigation, build capacity in the Victorian teaching profession in making consistent judgements by establishing a 3-year statewide initiative funded at an appropriate level to include the following elements:

- a. The appointment of three trained assessment support officers in each region, whose role would be to provide professional development and direct support to schools.*
- b. The provision, through the statewide/regional network of 3 days of professional development for all teachers in Years 1, 3, 5, 7 and 9 in 2008, and 3 days for all teachers in Years 2, 4, 6, 8 and 10 in 2009*
- c. The inclusion in the professional development days of activities designed to develop common understandings of the VELS, and on how to collect an appropriate body of evidence on which to base assessments*
- d. The inclusion in the professional development days of opportunities for all participating teachers to engage in moderation activities across schools.*
- e. Allocation of funding for an additional 2 student free days for all schools (matching the targeted year levels) for school level moderation, supported by regional officers.*
- f. The review, by the teams of regional assessment support officers of a random sample of assessments from approximately 20% of schools in each region in order to monitor consistency and accuracy and to validate teachers' assessments*
- g. The establishment, at the inception of the statewide assessment training initiative, of a well-defined process for reviewing the outcomes of the initiative*

Recommendation 3

It is recommended that the strengthening of assessment practices and making consistent judgements be included as key objectives within other system initiatives, for example, the performance and development culture in Victorian schools, and the improvement of pre-service teacher education programs.

Recommendation 4

It is recommended that current processes involving the comparison of aggregated VELS data with AIM data should continue, and that longitudinal evidence of the consistency

and accuracy of teachers' judgements be collected from an intensive study of a statewide sample of individual schools and individual teachers.

Recommendation 5

It is recommended that the relevance of assessment policies and practices be linked to the developing performance and development culture in Victorian schools.

Recommendation 6

It is recommended that a systematic program of moderation activities involving teachers meeting together to assess samples of student work be adopted for implementation within and across Victorian schools.

Recommendation 7

It is recommended that a comprehensive strategy for improving the consistency of teachers' judgements in Victorian schools be developed, and that this strategy should be informed and supported by a systematic evaluation of impact, benefits and costs.

1. INTRODUCTION

The Australian Council for Educational Research (ACER) was commissioned by the Department of Education in Victoria to undertake an investigation of current best-practice approaches to ensuring consistency of teacher judgements against P-10 state-wide standards.

The investigation encompassed national and international approaches that had been implemented or were under development. A key objective of the investigation was that the approaches identified should be based on current information, research and best practice, and that they should be supportive of the continued development of a learning and assessment culture in Victorian schools. The report of the investigation was required to provide key baseline information that would guide further work on the development and implementation of a coherent approach linking to and building upon the Victorian Essential Learning Standards (VELS), progression points, continua, assessment maps and other policy, advice and support materials.

Context

Approaches to developing consistency of teacher judgements is currently of major interest in relation to the implementation of curriculum and standards frameworks in Australian school systems. Standards-referenced assessment practices involve teachers in collecting a range of evidence of students' work, and making judgements about this work against state-based standards. It is widely recognised that formative assessment practices based on standards that provide maps of achievement in particular areas of learning are important in classrooms, and that they play an important role in improving learning. Teacher judgements against standards are also used for summative assessments, for example, in end of semester reports. In a variety of ways, teachers' judgements inform planning for future teaching, and reporting to parents and, in some jurisdictions, to the system.

Teachers' judgements against standards provide information that is used for a range of purposes, and it is essential that there is consistency in these judgements between teachers in the same school, across different schools, and over time. Approaches to developing such consistency are therefore seen to play a key role in strengthening assessment practices, and in increasing the validity and reliability of reporting against curriculum and standards frameworks.

In Victoria, consistency of teacher judgements is important in the context of the *Blueprint for Government Schools* and the Victorian Government's ongoing commitment to improving student learning outcomes. The goals and targets set for the education and training system as a whole, coupled with the reform agenda set out in the *Blueprint for Government Schools*, provide the context for the investigation. There are strong links between the approaches to consistency of teacher judgement, and the policy context set by the *Blueprint for Government Schools*, and the individual *Flagship Strategy Initiatives (FSI)*. Of particular relevance are the *FSI Student Learning initiatives (The Victorian Essential Learning Standards, Curriculum Planning Guidelines, The Principles of Learning and Teaching)* which, together with *Assessment and Reporting Guidelines and Advice* including the *Assessment Modules*, are designed to form an integrated, and coherent whole.

Findings about effective approaches to developing the consistency of teacher judgements will contribute to further work on the development and implementation of a coherent approach that will link to, and build upon, the *Victorian Essential Learning Standards (VELS)*, progression points, continua, assessment maps and other policy, advice and support materials that have already been developed. With the introduction of the VELS and the New Student Report Cards, teachers require support and advice to deliver consistent judgements on student achievement against the standards.

Objectives of the investigation

The key purpose of the project was to identify and outline approaches that, on the basis of available evidence, are most effective in improving the consistency of teacher judgements against state-wide P-10 standards. The specific objectives were to:

- Identify and outline approaches that, on the basis of available evidence, are most effective in improving the consistency of teacher judgements against state-wide P-10 standards
- Provide a detailed overview of each approach including any identified strengths/weaknesses, costs /benefits
- Document the rationale and principles underpinning the selection of the identified approach
- Provide a detailed overview of each approach, including an analysis of identified strengths and weaknesses and costs and benefits.

Improving consistency of teacher judgements

The issue of developing approaches to improve consistency of teacher judgements against state-wide standards is currently under discussion in all states and territories. In some states, jurisdictions have implemented large scale, system-wide initiatives related to improving consistency; in most states a range of resources focused on assessment practices is available to teachers in variety of formats; in other states plans for introducing new initiatives are afoot.

Recent research and practice relating to standards-referenced assessment and to formative assessment provides the basis for the development of resources, advice to teachers, and professional development activities, in various ways, across Australia. The issue of using data to improve learning has attracted increasing interest.

As curriculum and standards frameworks become embedded in school systems, appropriate ways of using these frameworks for curriculum planning, and for assessment and reporting are discussed in schools and at system-level. The development of technology now makes it possible for schools to report teacher judgements against standards to a central agency, and to receive information showing the trends in their own school's performance over time, and comparisons between the reported performances of students in their schools with the reported performances of students in similar schools. Data showing comparisons between the reported achievement levels of students against state-wide standards can be compared with the students' performance on state-wide tests at particular year levels.

In this context, the significance of teachers' judgements of their own students' achievement is now widely recognised. It is also recognised that the quality of these judgements is important, and that the consistency of judgements against a common standards framework might vary from school to school, unless there are common, shared understandings of the standards, and of the kind of evidence that indicates achievement against the standards. Thus there is an imperative for developing effective approaches to improving the consistency of teachers' judgements within schools, between schools, and over time.

Organisation of this report

The report begins with an outline of the methodology (Chapter 2) and a review of the research literature (Chapter 3). The review highlights major themes relating to the development and improvement of the consistency of teacher judgements. This is followed by an overview of current

practices in the Australian states and territories (Chapter 4), and of approaches in three international contexts (Chapter 5). The final section draws out the main findings of the investigation (Chapter 6), and makes some recommendations for building on current practices in the Victorian context, and the raft of initiatives that have been developed within the *Blueprint for Government Schools* (Chapter 7). A bibliography of all key references is included, along with the work cited in the report.

2. METHODOLOGY

The investigation was conducted between the end of October 2006 and February 2007. Data collection took place in November and December. Analysis of the data, and the drafting and revision of the report was done in January to April 2007.

The investigation was based on the premise that consistent teacher judgements are judgements made by teachers about students' performance in relation to state-wide standards. Such judgements are required to be a 'correct' reflection of students' achievements, consistent with judgements previously made for individual students, and with judgements made by other teachers for individuals or groups of students.

The investigation sought information about approaches in a variety of contexts. Teachers' capacity to make consistent judgements against state-wide standards is dependent on the development of shared understanding of the standards, that is, in the current Victorian context, the VELs and the related progression points. Further, consistency of teacher judgement is supported by collaborative planning of teaching and learning activities, and of assessment tasks. Consistency of judgement is achieved when teachers have opportunities to engage in moderation of samples of students' work, including the work of their own students and of other students.

Data and analysis

Several strategies were employed for the investigation:

- A search of data bases through the Cunningham Library at ACER
- Consultations: face to face, telephone interviews, email
- Desktop research of websites
- Review of relevant print and non-print resources
- Ongoing documentation of key resources
- Regular consultation with DoE Project Manager and the consultative group convened for the purposes of the project.

Selection of contexts for investigation

The investigation examined interstate and some overseas approaches. Data was gathered from all Australian government school systems.. State-wide standards had been set in almost all systems, were being developed, or reviewed.

In consultation with the DoE project manager, it was agreed that the international component of the investigation should focus on New Zealand, England and the province of Ontario, Canada. These contexts were selected for a range of reasons. England and New Zealand have nation-wide curriculum and assessment systems. In recent years there has been considerable interest in England in formative assessment, balancing processes of national assessment. Ontario was selected because of its relatively strong research base, and also because, like Victoria, it has a strong focus on school-level management and is part of a federal structure.

Consultations

Consultations were conducted with curriculum and assessment officers, through telephone and face-to-face interviews, in all states and territories. Some data was also collected from cluster leaders and schools. The extensive array of resources available to school leaders and teachers

through the websites of the jurisdictions were explored. Overall, the consultations identified approaches in a variety of contexts where emphasis is placed on the development of shared understanding of standards, collaborative planning and moderation.

Documentation

At the commencement of the project, a matrix was developed for the documentation of all information about different approaches. This matrix included the following elements:

- Contextual information about levels of schooling, and the nature and function of state-wide standards.
- Details of the evidence and research base of the approaches.
- Documentation of evidence of the achievement of consistency of teacher judgement, including examples of moderation activities.
- Costs and benefits of the different approaches.

Throughout the investigation, information about current approaches was mapped onto the matrix, and this enabled commonalities and differences between the reported approaches to be reviewed.

The identification of approaches that are cost-effective, that is, those which achieve particular outcomes with the least use of resources was a consideration in the investigation. As part of the examination of the identified approaches, the investigation collected evidence on the level of resources used under different approaches. Chief among these were teacher and project officer time, and the provision of professional development and support materials.

Selection of approaches

Approaches were identified as being effective in improving consistency of teacher judgements in terms of the extent to which they included the following elements, or particularly effective approaches to at least one of the elements:

- Teachers' knowledge and understanding of the standards
- Curriculum planning; opportunities for students to demonstrate achievement of standards
- Evidence of student achievement
- Assessment of the evidence against the standards
- Validation of teachers' judgements.

Evidence of improved consistency of teacher judgements

The consistency of the judgements teachers make against a set of standards lies in the extent to which the judgements correctly reflect student achievements, are consistent over time, and the judgements of teachers in different schools are consistent. This involves the following elements:

- judgements made for individual students are consistent with the judgements previously made for those students;
- judgements for all students in a range of contexts are consistent; and
- judgements made by teachers are consistent with judgements made for all students.

Professional learning

The strong links between the development of consistency of teacher judgements and teachers' professional learning were clearly evident in the investigation. The extent to which the identified approaches were aligned with research-based principles for teacher professional learning was therefore taken into account.

3. EVIDENCE AND RESEARCH BASE

Teacher judgements against state-wide standards raise issues relating to a number of educational practices. Teacher assessment is a central aspect of assessment and reporting practices, and in the context of this investigation, of standards-referenced assessment. Teacher assessment is an integral aspect of pedagogy, and part of the cycle of planning, teaching, assessment and review. The reliability and validity of teacher assessment is an issue involved in approaches to enhancing consistency of these assessments. Moderation is a specific assessment practice that provides one means of developing consistency, and involves considerable professional collaboration. Improvements in assessment practices, and in more consistent teacher assessments, grow from professional learning opportunities for teachers.

This overview of the research base is structured around four themes.

The first of these themes is the broad theme of assessment in education. There is a considerable body of recent research literature on assessment and assessment practices that has been influential in shaping approaches, and includes studies focused on the functions of assessment, on formative assessment, on teacher feedback, on classroom assessment, and on professional judgements. Of particular relevance to the consistency of teacher judgement has been the professional discussion of assessment for learning.

A second, more specific theme relates to standards-referenced assessment.

The third theme, moderation, focuses on processes involving contextualised teacher assessments, the role of these processes in achieving comparability of assessments against standards, and the reliability of reported results from these assessments.

The fourth theme of major relevance to this investigation concerns teacher professional learning, as a prominent factor in approaches to improving the consistency of teacher judgements.

Assessment

Teachers' judgements are interwoven into all assessment practices, and the quality of these judgements determines the extent to which assessment is valid and reliable. Assessment serves a number of critical functions, summarised by Black (1995, p. 14). The first of these functions is where assessment assists learning, wherein assessment information is used, by both teachers and pupils, to modify their work in order to make it more effective. A second function of assessment is in the certification of individual students, and the third function identified by Black concerns the public accountability of institutions. Any consideration of approaches to improving the consistency of teacher judgements must include all of these functions.

In the past decade, change and reform of assessment practices, both in the context of external standardised testing and in classroom assessments has been high on the agenda, in Australia and overseas. Curriculum and standards frameworks have been developed to provide a means of following students' progress in different learning areas throughout schooling. A snapshot of assessment in Australian schools presented by Cumming and Maxwell (2004, p. 106) noted that:

A range of assessment practices that are believed to enhance student educational development and pathways are firmly in place in Australia. The rhetoric, as in all education, still exceeds the practice. The quality of assessment practice is uneven, most notably in the compulsory years of schooling where collaborative moderation and professional development to explore common understandings of standards and expectations are still developing. Quality still seem to be related

to teacher experience (both too limited and too entrenched), familiarity with assessment issues, and resources.

One of the major shifts in emphasis in recent years has been toward greater interest in the interactions between assessment and classroom learning (Hattie & Jaeger, 1998).

Assessment and Classroom Learning

A number of factors have contributed to this shift in emphasis. A highly influential review of relevant research literature on formative assessment (Black & Wiliam, 1998) had led to significant discussion of formative assessment. The review covered a wide range of published research and provided evidence that formative assessment raises standards, and that current practices are weak. (Black *et al.*, 2003).

Black and Wiliam have focused on ways of helping teachers to put the research findings into action¹. Their work (Black *et al.*, 2003; Black, Harrison, Lee, Marshall, & Wiliam, 2004; Black & Wiliam, 1998) is widely disseminated and has contributed to debates on testing and assessment. One aspect of the debate, particularly in England, has been on the need to balance the focus on external, standardised testing of recent years with a focus on classroom assessment. It is well understood that data from standardised testing provides teachers with valuable information that can be used in many ways to improve learning, but equally, that formative assessment in classrooms plays a crucial role in improving learning.

The importance accorded to assessment in learning is reflected in a major study in Scotland, the *Assessment is for Learning Programme* (2003-2004), which was designed to:

Bring together the various purposes of assessment into a single coherent framework which would answer questions accountability, standards and monitoring of progress and performance but which emphasised the role of assessment in supporting individual pupils' learning in the classroom... [In the Programme] the 'big ideas' about assessment are that learners learn best when

- *they understand clearly what they are trying to learn, and what is expected of them*
- *they are given feedback about the quality of their work, and what they can do to make it better*
- *they are given advice about how to go about making improvements*
- *they are fully involved in deciding what needs to be done next, and who can give them help if they need it* (Condie, Livingstone, & Seagraves, 2005, pp. 2-3)

Increased emphasis on classroom assessment has also stemmed from the range of educational reforms. Earl's book for teachers, *Assessment as Learning* (Earl, 2003) draws on the body of research about classroom assessment. She cites Leithwood's analysis of the characteristics of large-scale reform that resonate with government initiatives in Australia and other countries:

- *Vision and goals*

¹ The website of the Assessment Reform Group draws this work together:
<http://www.assessment-reform-group.org>

- *Standards*
- *Curriculum frameworks and other teaching resources*
- *Focus on teaching and learning*
- *Accountability and incentives based on performance*
- *Coherent and integrated policies*
- *Sufficient funding and workable governance structures.* (Earl, 2003)

Several of these characteristics revolve around teaching, learning, and assessment. The embedding of standards and curriculum frameworks has highlighted ways that teachers are constantly involved in assessment, both formative and summative. While the work of the Assessment Reform Group has been focused on formative assessment, and how teachers can be supported in effective formative assessment, teachers also undertake summative assessments, which are reported to various audiences. Harlen (2005b, p. 247) defined teachers' summative assessments as:

The process by which teachers gather evidence in a planned and systematic way in order to draw inferences about their students' learning, based on their professional judgements and to report at a particular time on their students' achievements.

This is useful means of pinpointing what teachers do when they make judgements against state-wide standards and report these judgements to multiple audiences. It is clearly essential that the professional judgements made in this process are valid and reliable, comparable and consistent, and that efforts to improve consistency should be a high priority.

Developing understanding of the importance of formative assessment also means that it is important to strengthen teachers' practices in formative assessment. Since 2002, the OECD Centre for Educational Research and Innovation (CERI) has examined formative assessment. Studies of practices in eight countries (including Australia) and literature reviews have clarified the concept of formative assessment and how policies supporting the use of formative assessment can be developed. A key finding from the studies is that improving formative assessment raises student achievement:

In classrooms, formative assessment refers to frequent, interactive assessments of student progress and understanding to identify learning needs and adjust teaching appropriately. Teachers using formative assessment approaches and techniques are better prepared to meet diverse students' needs – through differentiation and adaptation of teaching to raise levels of student achievement and to achieve greater equity of student outcomes. (OECD, 2005).

The phrase 'assessment for learning', as Black *et al.* (2003) point out, has become a common substitute for 'formative assessment', yet there is possible ambiguity in this label.

Others have linked assessment for learning with pedagogy:

What we have called the spirit of Assessment for Learning is instantiated in the way teachers conceptualise and sequence the tasks undertaken by pupils in the lesson. The nature of these tasks affects all subsequent interactions within the class. (Marshall & Drummond, 2006, p. 147).

Formative assessment is also connected to the concept of monitoring progress against standards. In an exploration of formative assessment and what she calls 'the conundrum of progression in English', Marshall (2004, p. 112) concludes that:

What formative assessment does provide is a kind of framework for thinking about progression even if it cannot be neatly described.

There is an extensive body of research evidence about both formative and summative assessment. The purpose of the brief overview of some key aspects here has been to point to the evidence base for arguing that efforts to improve the consistency of teachers judgements should be a priority, and that there is a “critical need for sustained and cooperative efforts on changing classroom assessment practices” (Tierney, 2006, p. 260).

In the literature, there is significant recognition of the importance of teacher assessments. For example, in a discussion of performance assessment, Ryan argues that more consideration needs to be given to

... teachers who are daily front-line observers, assessors, and evaluators. Teachers are well-positioned to comment on performance given their long-term exposure to the evidence produced by the student. (Ryan, 2006, p. 103)

In a systematic review of research on the reliability and validity of teachers’ assessment for summative purposes, Harlen (2004; 2005a) found evidence of low reliability of teachers’ judgement made in certain circumstances, but also identified ways of overcoming these deficiencies.

The systematic review conducted by the Evidence for Policy and Practice Information and Coordinating Centre (EPPI-Centre) at the Institute of Education, University of London (Harlen, 2004) was prompted by the work of the Assessment Reform Group in the UK. The research carried out by Black and Wiliam (1998) indicated that assessment for formative purposes benefits teaching and learning, and raises standards of student performance. In this context,

Assessment by teachers has the potential for providing summative information about students’ achievement since teachers can build up a picture of students’ attainments across the full range of activities and goals, Although assessment by teachers is used as the main source of information in some national and state assessment systems, in other countries, it has the image of being unreliable and subject to bias. This review was undertaken to provide some research evidence about the dependability of summative assessment by teachers and the conditions which affect it. (Harlen, 2004, p 1).

The methodology for the review followed procedures developed by the EPPI-Centre, and defined criteria for guiding a search for studies that dealt with some form of summative assessment conducted by teachers, involving students aged from 4- 18 years. The search yielded 431 papers, from which 30 studies were finally selected.

For the purposes of the systematic review, assessment by teachers for summative purposes was defined as

any activity in which teachers gather evidence in a planned and systematic way about their students’ learning to draw inferences based on their professional judgement to report achievement at a particular time. (Harlen, 2004, p 1).

The review concluded that ‘for dependable summative assessment (that is, with construct validity protected and optimum reliability) the requirements are as follows’:

- *Decisions about the domain of knowledge, skills and the attributes of learning to be assessed that are justified in terms of how learning takes place;*
- *A valid sample of student behaviour in the domain;*
- *Criteria for judging the sample that are well matched to the goals of the work, of the curriculum and the domain;*
- *Procedures for the reliable and unbiased application of the criteria;*
- *Procedures for reporting and communicating with users of the assessment outcomes. (Harlen, 2004, p 85)*

The review has clear implications for policy and practice. Overall,

Solutions to the problems of inconsistency in the type of evidence used and in the application of criteria suggested by the studies focused on five types of action, relating to: the specification of the tasks, the specification of the criteria, training, moderation, and the development of an 'assessment community' within the school allied to increased confidence in the profession's judgement of teachers. (Harlen, 2004, p 96).

Implication for policy and practice were set out, and several of these have clear relevance to the present investigation:

Policy

- *There is a need for resources to be put into identifying detailed criteria that are linked to learning goals, not specially designed assessment tasks. This will support teachers' understanding of the learning goals and may make it possible to equate the curriculum with the assessment tasks.*
- *It is important to provide professional development for teachers in undertaking assessment for different purposes that address the known shortcomings of teacher assessment.*
- *The process of moderation should be seen as an important means of developing teachers' understanding of learning goals and related assessment criteria.*

Practice

- *Teachers should not judge the accuracy of their assessments by how far they correspond with test results but by how far they reflect the learning goals.*
- *There should be wider recognition that clarity about learning goals is needed for dependable assessment by teachers.*
- *Teachers should be made aware of the sources of bias in their assessments, including the 'halo' effect, and school assessment procedures should include steps that guard against such unfairness.*

- *Schools should take action to ensure that the benefits of improving the dependability of the assessment by teachers is sustained (e.g. by protecting time for planning, assessment, in-school moderation etc).*
- *Schools should develop an ‘assessment culture’ in which assessment is discussed constructively and positively and not seen as a necessary chore (or evil.) (Harlen 2004, pp 96-97.)*

Standards and standards-referenced assessment

A recent international review of curriculum and assessment frameworks focused on The Arts indicated that most countries use assessment approaches that require teachers to use professional judgement against a standard, and provide a variety of materials to help teachers with these assessments (Taggart, Whitby, & Sharp, 2004).

Sadler’s seminal 1987 paper on standards-referenced assessment highlighted the centrality of teachers’ qualitative judgements. He described standards as ‘fixed points of reference for assessing individual students’.

Standards-referenced assessment draws upon the professional ability of competent teachers to make sound qualitative judgements of the kind they make constantly in teaching. The justification for putting effort into the development of a standards-referenced assessment system lies in the belief that teachers’ qualitative judgements can be made dependable provided that (a) standards are developed and promulgated in appropriate forms, and (b) teachers are given the relevant conceptual tools and practical training. (Sadler, 1987)

The prevalence of the now established need for teachers to make professional judgements against standards leads to the search for processes that will build the dependability of these judgements, and moderation is a process frequently identified as a key means of improving teacher judgements.

Moderation

Moderation is widely seen to play a critical role in improving consistency of teacher judgements. A number of perspectives illustrate this.

In his discussion paper on assessment and reporting for the Queensland School Curriculum Council, Maxwell argued that moderation is desirable at all levels of the education system. He referred to it as

moderation for improvement in the sense that it is directed at enhancing the quality of schools’ assessment programs and teachers’ judgements of student progress. In other words, this form of moderation is explicitly linked to professional development and directed at overall improvement in the education system. (Maxwell, 2002)

In the *Consistency in Classroom Assessment: Support Materials for Educators* compiled by the Council of Ontario Directors of Education (2006) it is noted that

In the literature, the most common activities to improve assessment are moderation. Moderation can be defined as activities that allow teachers to compare their judgements with other teachers in their school. ... In addition to moderation activities, the research cites activities in which teachers

collaboratively plan their teaching and assessment as a powerful strategy to improve consistency.

Harlen links the improvement of teachers' assessments with professional development and quality assurance and quality control:

In our view it is possible to enhance the quality of teachers' assessments through moderation processes that support professional development. By doing so we would achieve assessment results which would give dependable information about pupils' and students' performance across the wide range of aims of education. (Harlen, 1994)

The moderation of teachers' judgement, necessary for external uses of summative assessment, can be conducted so that it not only serves a quality control function, but also has an impact on the process of assessment by teachers, having a quality assurance function as well. (Harlen, 2005b).

The process of moderation should be seen as an important means of developing teachers' understanding of learning goals and related assessment criteria. (Harlen, 2005b).

The consistency and comparability of teacher judgement is a constant theme. Here, the importance of common understandings is emphasised:

Developing a common understanding of syllabus outcomes is an important step in working towards achieving consistency of a teacher's judgement and comparability between the judgements of teachers. (Bruniges, 2001)

Moderation involves teachers working together to look at samples of student work, and the deep value of this is recognised here:

The value of looking at student work resides in its potential for bringing students more consistently into deliberations among teachers. Looking at students' work has the potential to expand teachers' opportunity to learn, to cultivate a professional community that is both willing and able to inquire into practice, and to focus school-based teacher conversations directly on the improvement of teaching and learning. (Littie et al., 2003, p. 192)

Moderation practices are found in many contexts. For example, in recent work in Singapore on developing classroom assessment tasks aligned with authentic intellectual standards, brought about by the recognition of the *importance of alternative or authentic assessments such as project work, portfolios, major research paper, performance based assessment*, teacher moderation of actual assessment and student work from their schools was conducted. (Koh et al., 2006)

There is a considerable body of evidence about the effective management of moderation processes, for example, the 1998 DETYA study identified several factors that made moderation difficult:

- *Where students' work showed insufficient evidence*
- *Where teachers set different tasks which needed clarification of what had been taught and how*
- *Where tasks were poorly set*
- *When a piece of work showed achievement of outcomes at more than one level*

- *Where teachers had difficulty in interpreting the exact meaning of the CSF statements. (Consistency Project Report. Enhancing Consistency in Teacher Judgement of Student Work. Victorian report., 1998)*

Evidence such as this highlights the connections between assessment, moderation, and teachers' professional learning.

Teacher professional learning

Loucks-Horsley's work on designing professional development for teachers of mathematics and science identifies a number of effective strategies for professional learning. One of these is the collaborative examination of student work. The assumptions about learning, teaching and professional development that underlie this strategy are that:

Student learning is the ultimate outcome for professional development, and the closer the professional development opportunity brings teachers to student learning the better, this is different from professional development that focuses on teaching practices. Examining student work and student assessments focuses teachers' attention on the consequences of their teaching for learners, which demonstrates to teachers discrepancies between what they believed they were teaching and what students appear to have learned. (Loucks-Horsley et al, 1998)

Harlen, in reviewing the research on the reliability of teachers' summative assessments argued that, in order to increase the validity of these assessments,

Professional development is need for a variety of purposes: to ensure that teachers have the skills, knowledge and support to conduct assessment effectively; to ensure quality assurance and quality control so that users can have confidence in teachers' judgements; and to help teachers reconcile the dual role they are required to take in both promoting and judging learning. (Harlen, 2005b)

Professional development is needed for a variety of purposes: to ensure that teachers have the skills, knowledge and support to conduct assessment effectively; to ensure quality assurance and quality control so that users can have confidence in the teachers' judgements. (Harlen, 2004)

A key finding on the role of moderation from the Queensland New Basics project was that

Teachers' participation in various stages of moderation was one of the most important contributors to professional skills enhancement and to developing confidence in applying the model for grading students' Rich Tasks performance. (Matters, 2006)

Gipps, Clarke & McCallum (1998) reported that teachers need much more help in assessment processes which can best be provided in the form of exemplification materials and some form of group moderation.

In a discussion paper for the Queensland School Curriculum Council, Maxwell (2002) argues that moderation can serve two important purposes: moderation for accountability and moderation for improvement. He describes these two forms of moderation as follows:

- *Moderation for accountability provides official confirmation of assessments that are reported or used publicly, whether for individuals or cohorts of students, it therefore involves some external control mechanism or validation requirement.*
- *Moderation for improvement develops the capability of teachers to make consistent and comparable judgements but lacks the external control mechanism or validation requirement of moderation for accountability; it therefore involves collaborative processes supporting the professional development of teachers.* (Maxwell, 2002, p. 1).

Maxwell also makes the link to professional development, as a further benefit of moderation.

Moderation for improvement is essential for developing coherence and consistency across the educational system. It also offers the most powerful form of professional development. (Maxwell, 2002, p. 21)

Assessment systems

Several studies present evidence to suggest that there are some essential practices that need to be in place to ensure the quality of teacher judgements against standards.

Wiggins (1989) suggests that a truly authentic/innovative assessment system should meet the following five criteria:

(1) criterion/standards referenced, (2) formative, (3) moderated, (4) clear in the progression of educational development, and (5) a substantive assessment framework that describes the achievement variables that are valued and thus worth assessing.

In the systematic review of teachers' summative assessments, Harlen (2005b) concluded that improving the dependability of these assessments focuses on five kinds of action identified in the studies:

... the specification of the criteria for assessment; the training of teachers in assessment methods; moderation of the outcomes of teachers' judgements; and the development of an 'assessment community' within each school designed to increase confidence in the professional judgement among teachers themselves as well as among users of assessment. (Harlen, 2005b)

These have implications for policy, practice and research, including Harlen argues, moderation:

the process of moderation should be seen as an important means of developing teachers' understanding of learning goals and related assessment criteria. (Harlen, 2005b)

A further set of criteria can be found in the Queensland New Basics Research Report (2004). Drawing on unpublished work by Allen (2003) the report identifies five elements of an effective assessment system 'as underlying the best of a wide range of assessment policies and practices'. The usefulness of these five elements was demonstrated in a conference paper presented by Matters (2004).

1. *What (if any) guidelines must teachers/schools use in planning?*
2. *What (if any) formal plans for student learning and achievement must teachers/schools make?*

3. *What (if any) evidence of student achievement must be produced?*
4. *How (if at all) is this evidence assessed against the guidelines and plans?*
5. *How (if at all) are teacher judgements of student achievement validated?*

In the descriptions of approaches to developing consistency of teacher judgements in Australia, New Zealand, England and Canada that follow, the elements identified in the New Basics Report were used to organise the descriptions so that the effectiveness of the approaches could be identified. The investigation of current practices set out to identify ‘best-practice’ approaches, and this framework provided a means for analysing the different approaches. This framework made it possible to analyse the coherence and comprehensiveness of the different approaches, and to look closely at particular components, such as moderation, within these systems.

4. CURRENT PRACTICES IN AUSTRALIAN STATES AND TERRITORIES

This chapter describes and analyses the range of approaches designed to develop consistency of teacher judgement identified in the investigation of practices current within Australian school systems in 2006.

Table 1 provides a map of these approaches, indicating the curriculum and standards frameworks currently set for states and territories, or under development. The key components of each approach are listed in Table 1, and these are described in more detail for each state and territory.

The investigation revealed some common aspects in approaches designed to improve the consistency of teacher judgements. In order to draw out the commonalities and differences, and the relative effectiveness of these initiatives and approaches, five key elements of an effective assessment system underlying a wide range of assessment policies and practices reported in the research literature (Matters, 2004, p. 14) have been used as organisers for the descriptions. These organisers are as follows:

- Teachers' knowledge and understanding of the standards
- Curriculum planning; opportunities for students to demonstrate achievement of standards
- Evidence of student achievement
- Assessment of the evidence against the standards
- Validation of teachers' judgements.

Overall, evidence of these elements was found, to varying degrees, in all of the approaches investigated, including a variety of print and web-based resources to support teachers' use of the standards, and a range of professional learning and moderation activities varying from in-school to state-wide activities. In general though, there appears to be little documentation of the implementation of these elements, or of their benefits and costs.

The account of Victorian approaches is presented first, followed by the other states and territories in alphabetical order.

Table 1: Approaches to consistency of teacher judgements against state-wide standards in Australian states and territories

	Standards Framework	Key Components of the Approach
Victoria	<ul style="list-style-type: none"> Victorian Essential Learning Standards (2005) 	<ul style="list-style-type: none"> VCAA progression points VCAA assessment maps VCAA sample assessment tasks (in development) DoE Sample multi-domain tasks DoE English and mathematics developmental continua Professional learning Consistency of Teacher Judgement CD-ROM VCAA annotated work samples (in development)
Australian Capital Territory	<ul style="list-style-type: none"> Every Chance to Learn: Curriculum framework for ACT schools P-10 * 	<ul style="list-style-type: none"> Workshops outlining A-E Reporting Some school based moderation
New South Wales	<ul style="list-style-type: none"> K-12 Subject Syllabuses Primary Curriculum Foundation Statements 	<ul style="list-style-type: none"> Consistent Teacher Judgement Online Professional Learning Modules Consistent Teacher Judgement in Action Consistent Teacher Judgement website http://www.curriculumsupport.education.nsw.gov.au/consistent_teacher/index.htm Assessment Resource Centre http://arc.boardofstudies.nsw.edu.au/
Northern Territory	<ul style="list-style-type: none"> Northern Territory Curriculum Framework 	<ul style="list-style-type: none"> Cluster moderation activities Building Better Schools: Standards Validation Project Curriculum eTool (CeTool) Multilevel Assessment Program writing moderation System annotated samples Assessment for Learning Project

Queensland	<ul style="list-style-type: none"> Queensland Curriculum, Assessment and Reporting Framework * 	<ul style="list-style-type: none"> New Basics Trial moderation New Basics Rich Tasks http://education.qld.gov.au/corporate/newbasics/html/richtasks/richtasks.html State-wide assessments to be introduced Assessment Bank in development
South Australia	<ul style="list-style-type: none"> South Australian Curriculum Standards and Accountability Framework 	<ul style="list-style-type: none"> SACSA Team & Standards Team Consistency of Teacher Judgement CD-ROM Moving Forward with SACSA: Resources for leaders to support the understanding and practice of consistency in teacher judgement CD-ROM SACSA Moderation Protocol Professional learning Let's Talk Assessment ... newsletter
Tasmania	<ul style="list-style-type: none"> Tasmanian Essential Learnings 	<ul style="list-style-type: none"> Quality Moderation of Assessment Process (QMAP) http://wwwfp.education.tas.gov.au/oer/Qmap/index.htm QMAP support team 33 Assessment Support Leaders State-wide moderation days Guiding assessment tasks Learning, Teaching and Assessment Guide website http://www.ltag.education.tas.gov.au/ Moderation protocol
Western Australia	<ul style="list-style-type: none"> Outcomes and Standards Framework 	<ul style="list-style-type: none"> Making Consistent Judgements professional learning days Western Australian moderation protocol Literacy and numeracy assessment exemplars and profiles

* In development

VICTORIA

Context

In Victoria, until 2005, the Curriculum and Standards Framework and the revised CSFII provided the standards against which teachers reported students' progress. The Victorian Essential Learning Standards (VELS) replaced CSF11 as the curriculum and reporting framework in 2006. Consistency with the CSF was maintained in terms of the broad function and structure of standards (six achievement levels from P-10) but in addition to the standards for the disciplinary domains, two new strands were introduced, Physical, Personal and Social Learning and Interdisciplinary Learning. Furthermore, in tandem with the development of the standards, other key resources such as the Principles of Learning and Teaching and new student report cards were developed to provide a coherent and interlocking approach to curriculum design, teaching and learning, assessment and reporting.

The VELS are organised in strands, domains and dimensions with learning focus statements for each domain and standards for each dimension. The standards define what students should know and be able to do at different levels of schooling. The standards provided at each of the six levels are broadly equivalent to Prep and Years 2, 4, 6, 8 and 10. The standards are set at a challenging level of competence and are intended to stretch students in their learning.

Information about student progress against the standards is generated through teacher judgement using teacher designed tasks, teacher judgement against externally designed tasks, through externally designed and assessed measures (state-wide AIM testing of English and Maths in Years 3, 5, 7 and 9), national and international sample-based testing (Science, ICT, Civics and Citizenship, PISA and TIMSS), and any other external tools schools may choose to use such as DART or TORCH.

In Victoria approaches to ensuring consistency are determined and provided jointly by the VCAA and the various school sectors, Government, Catholic and Independent. Historically, approaches have focused predominantly on providing sample material, for example assessment maps, sample tasks, marking guides and annotated student responses, to use at the point of, or more usually, post assessment and at the discretion of individual teachers, faculties, schools and regions. The focus of these materials often tends to be on localised, post hoc, consensus moderation.

Approaches to Consistency

Supporting knowledge of the standards

The introduction of the VELS in 2006 has meant that significant work has been done and continues to be done to support teachers' understanding of the new Standards. The VCAA has developed Progression Points, descriptors that act as indicators of progress towards the standard for the level. Teachers use standards and progression points together with assessment maps to make on-balance judgements about student performance for informal and formal monitoring and reporting of student achievement.

Advice on each domain, to be published by VCAA in 2007 and focussing on relationships with other domains and with curriculum at Years 11 and 12, is intended to continue to build teacher knowledge and understanding of the standards.

The Catholic Education Office, Melbourne has delivered professional development on understanding the VELS for curriculum coordinators and principals of primary schools. The training comprised two workshops. One of the professional development modules, Assessment & Moderation Workshop 2, focussed on assessment and the consistency of teacher judgements. This has been distributed on CD-ROM to Catholic primary schools. There is no evidence available as

to how this resource is being used in schools and there has been no similar support for secondary teachers in the Catholic sector.

Curriculum planning

There is very little material from Victoria about planning for learning that incorporates principles of making consistent judgements. However, planning to provide students with opportunities to demonstrate achievement of the standards is embedded in the work of schools and clusters where teachers are able to collaborate in developing jointly planned assessment activities. Additionally, professional learning days related to the implementation of VELs incorporated a focus on unit planning using backwards design, that is, planning from the standards and outcomes in the VELs that students are expected to achieve. Combining professional learning which builds understanding of the curriculum with planning assessment for learning is a beneficial approach, however the cost of providing teachers with extensive professional learning time is necessarily high.

A useful resource for developing consistency in teacher judgement is the Consistency of Teacher Judgement CD-ROM developed in a project of the same name in 2000. The content of the CD emphasises the importance of jointly planning outcomes based units of work with colleagues. A step-by-step process for doing this is outlined on the CD which was made available to all schools at the time it was produced. New copies of this material are not available and anecdotal evidence suggests that few schools continue to make use of this material. (It is interesting to note that this resource is extensively cited, for example, in a document compiled by the Council of Ontario Directors of Education (CODE, 2006).

Evidence of achievement

Teachers make judgements about student achievement of the Standards against a body of evidence of students' work. Teacher's ability to select work samples that appropriately demonstrate achievement of the Standards is crucial to improving consistency when making on-balance judgements. In Victoria, the DoE and VCAA assessment advice highlights the importance of assessment as an ongoing process of gathering, analysing and reflecting on evidence to make informed and consistent judgements to improve future student learning, and most importantly choosing assessment instruments and processes which measure what they are intended to measure. If this is the case then the body of evidence from which judgements about achievement are made is likely to be of good quality and provide a better basis for consistency.

Advice and statements about what is required to build a good body of evidence set the context, and resources demonstrating what effective assessment tasks look like provide a practical way of building consistency through assessing a quality body of evidence of student achievement. Sample assessment tasks provide just such a practical approach to improving the body of evidence on which teachers base their judgements. The VCAA is developing a suite of sample assessment tasks incorporating learning sequences, assessment criteria and assessment guides available to teachers through the VCAA website. A similar resource developed by the DoE is available online and focuses on multi-domain tasks as a model for the process of developing rich assessment

The English and Mathematics Developmental Continua support an evidence-based approach to assessing progress. Online sample tasks help locate students on the continua by linking tasks the standards and progression points in the VELs. Anecdotal evidence from schools suggests that the Continua are being used by teachers to assess their student's achievement and are valuable. However, there is no evidence available of whether this has yet led to greater consistency of judgement.

Assessment of evidence against the standards

In 2001 the VCAA published a kit of annotated work samples for each key learning area which were designed to assist teachers to develop a common understanding of student achievement against the standards of CSFII. The kit was accompanied by a state-wide professional learning program to assist teachers in using the samples to improve their assessment of student performance. The professional learning activities were designed to reach all regions and KLA leaders in schools, however the materials are no longer accessible on website or other media.

Currently, VCAA Assessment maps are being developed and will provide annotated work samples for each standard and progression point in each dimension of each domain. The Assessment maps currently provide samples of student work at the standard and have been carefully aligned to the VELs, clearly illustrating work which is at or progressing towards the standard. As such they provide a good benchmark of work at the standard and are useful for building consistency in that teachers can use them as references when assessing their own students' work or to build their own collection of samples.

Validation of teachers' judgements

There is no single strategy for validating teachers' judgements in place in Victoria. Moderation is frequently suggested in assessment support materials provided by the DoE and the VCAA but methods of moderation and moderation advice are more difficult to find.

There is evidence that some schools and clusters engage in a form of "social moderation"². One cluster has developed professional learning programs based on middle years literacy, numeracy, thinking and engagement. Network meetings are held once a term with up to 15 schools including the five key cluster schools in attendance. For literacy teachers, common assessment tasks were set for writing across the cluster and participating teachers worked together to reach consistent judgements in their assessment of these tasks. An effective cycle of moderation meetings has been established with up to four meetings per year. Despite this there is no conclusive evidence of improved consistency in teacher judgements.

Some schools have developed their own systems of moderation across year levels, and have also moved towards greater consistency in the use of assessment tools which have been changed to better match the VELs standards. One school, for example, has indicated that their own data shows that teachers have raised their expectations for student achievement and ongoing monitoring and assessment has assisted improved literacy levels, particularly in Years P-2 across the school over time.

The lack of a system-wide approach to moderation and a supported method of moderation means that there is little data available on the efficacy of the various systems of moderation in schools and clusters.

Targeted Resources and Professional Learning

A range of professional learning opportunities have been provided in Victoria to support consistency of teacher judgement. The Consistency of Teacher Judgement Project was a state-wide initiative in which all schools were invited to participate and for which a suite of resources

² In the *New Basics Research Report* (2004) social moderation is defined as: the process of validating teacher judgements of the standard of student work by having those teachers' judgements reviewed by their peers, internally within the same school and/or externally from a different school. ... Consistent and valid moderation possess three essential features: standards (explicit and a priori); evidence (student work); and consensus (between teachers judging standards of student work) (p. 164).

was produced in 2000, including a training and development CD-ROM. The materials were produced by the South Australian Department of Education, Training and Employment in collaboration with the Victorian Department of Education, Employment and Training and the Queensland School Curriculum Council. The project was supported by funding from the Commonwealth Department of Education, Training and Youth Affairs. The CD-ROM includes four strategies to assist teachers develop consistency in judging student work:

- Elaborating outcome statements
- Common assessment tasks
- Moderation of student work samples
- Collaborative planning of outcomes based units of work.

Focus questions and a step-by-step process were provided for each of the above. A state-wide curriculum day was devoted to developing consistent teacher judgements in October 2000. The CD-ROM is still available in schools, but new copies are not available and anecdotal evidence suggests that few schools continue to make use of this material.

More recently, the Assessment Professional Learning Modules published on the Department of Education website in 2006 are designed to enhance the capacity of teachers in assessing student performance. Module 5 focuses specifically on consistent judgement through planning, moderation and protocols. It includes a very brief but comprehensive outline of ways of planning for ongoing consistency in assessment processes, but the focus is primarily on localised, unchecked, consensus moderation processes. The Department has also developed a sequence of two workshops based around PowerPoint presentations which school leaders can deliver to teachers which provide advice about the approach to making judgements against the standards and progression points and using progress maps. A moderation exercise drawn from the Assessment Professional Learning Module 5 is included. Some schools have reported using the professional learning modules but this resource is not being used everywhere and there is no evidence of how the modules are impacting on teacher judgements. There is no evidence of system-level scrutiny of how schools implement this professional learning, or if indeed they use it at all.

While not focussing specifically on developing teacher capacity to make consistent judgements, some regions have organised a number of professional development activities on progression points, assessment maps, the continua and the standards with a key focus on familiarisation. Additionally, individual schools and clusters have developed their own approaches to consistency.

The Assessment for Learning website was developed by Curriculum Corporation on behalf of the education departments of the States, Territories and Commonwealth of Australia. It contains an “Assessment for Learning” site which provides assessment rubrics, annotated work samples and assessment advice to support teachers in making consistent judgements. The Department of Education Science and Training also provides links to assessment practices in the other states. There is no available evidence of the extent to which teachers are aware of these resources or use them in any structured way in Victorian schools.

Evidence of Improved Consistency

The introduction of A-E reporting has sharpened the focus on accurate and consistent assessments against the standards. Anecdotal evidence suggested that many schools were unconcerned with the accuracy or consistency of teacher judgement against the CSF, however the A-E reporting makes direct links with the VELS standards and has become an impetus for prioritising consistency and accuracy in assessment. While awareness of consistency and reliability is being established in the light of A-E reporting there is no evidence as yet of improved consistency.

A comparison of teacher judgement data with student performance data in the Achievement Improvement Monitor (AIM) for English and Mathematics shows that for the cohort within each year group, student achievements in the AIM are spread over one or more VELs levels whilst teachers' own assessments fall within a relatively narrow band, usually corresponding to a single VELs level. While revealing a high level of consistency over time, it appears that teacher judgement data substantially understates the proportion of students who are achieving well above and well below the expected level. Standardised test data from the AIM indicates that teachers are not looking beyond the expected standards for the year levels in which they teach. This would indicate that teachers' understanding of how to use the Standards is a weak point and that support for using the Standards throughout the assessment process is required.

Teachers assess student writing as part of the AIM using given criteria and mark schemes and pre-marked samples of student work. The VCCA receives only a small number of enquiries from teachers who have noticed a considerable discrepancy between their own internally assigned scores and external AIM scores. The VCAA sees this as evidence that there is a high correlation between the teacher-assessed and centrally-assessed scores. This, and anecdotal feedback from teachers, suggests that the process used in developing common assessment criteria and publishing annotated work samples is useful in ensuring consistent marking. However, this is a common task with detailed marking guides and support, which is not the case for general class assessments.

Summary

Victorian schools have been well resourced in terms of materials that support consistent teacher judgement. A strong suite of resources which includes assessment advice, annotated work samples, sample assessment tasks and professional development modules is all available to teachers and schools. However there is no requirement that the resources be used or if they are, of common approaches in the way that they are used. These resources are readily accessible given that the majority of them are online through the DoE and VCAA websites; however there is no evidence of the extent to which they are used.

The introduction of the VELs has been supported by a strong program of professional learning and advice from the DoE and VCAA and anecdotal evidence from schools suggests that teachers have good knowledge of the standards and progression points.

Approaches to validating teacher judgements in the form of comprehensive moderation programs supported by advice, professional learning and in some cases mandated programs and activities organised through central or cluster bodies are not well developed in Victoria. Moderation is embedded in professional learning modules and has been a focus in the past but might be highlighted and supported more explicitly to help build consistency of teacher judgement state-wide.

* * * * *

AUSTRALIAN CAPITAL TERRITORY

Context

Currently in the ACT there is no state-wide standards framework, in line with the policy of school-based curriculum development. However, a curriculum renewal process is being implemented. Every Chance to Learn - Curriculum framework for ACT schools P-10 is currently in the third phase of development. Due to be completed and published by late 2007 and implemented in 2008, the curriculum framework will outline the curriculum for students in all ACT schools, preschool to Year 10. The framework is sequenced in four bands of development: early childhood, later

childhood, early adolescence and later adolescence. Teachers will make judgements about where students are at the end of each of these bands in Year 2, Year 5, Year 8 and Year 10.

Markers of progress indicate the typical achievement expected of students at the end of each band. It is acknowledged that most students will make steady progress as they move through schools, however provision for exceptional progress where students may reach of markers of progress ahead of the end of the band of development or be slower to demonstrate achievement exists within the curriculum framework draft. As part of the curriculum framework development process, a final decision has not been made in relation to reporting against the Essential Learnings at the different stages identified within the document. There will be ongoing support for teachers in assessment and consistency of judgement, but the form of this support remains unclear. It may occur at a system level or if identified at a school level delivered by departmental staff or outside bodies; the details of such a program have not been finalised.

Approaches to Consistency

Supporting knowledge of the standards

With the move to introduce territory-wide standards in the ACT, centrally organised workshops have been held to introduce teachers to the new curriculum framework. This has often been done in conjunction with training on the A-E reporting system. The efficacy of this training will not be apparent until the standards are fully adopted and in use in ACT schools.

Assessment of evidence against the standards

The introduction of A-E grading in ACT schools has been supported by workshops that address what work at each grade level might look like. This was followed up with further workshops which looked at how four schools had managed the assignment of grades and identified successful approaches. The introduction of A-E reporting has clearly identified that schools need to moderate judgements more effectively. Feedback has suggested that in primary schools in particular this is one of the benefits of the introduction of A-E reporting.

Validation of teachers' judgements

Validation of teachers' judgements, another key element, has been recognised as an important factor in the assessing and reporting process, but is only used in the senior secondary context with any regularity or rigour at present.

Moderation occurs in senior secondary schools at Years 11 and 12 where agreed curriculum documents are in place. This moderation takes the form of a structured, consensus-based peer review of portfolios of student work. Moderation takes place twice yearly in March and August on a specific day in each subject. Teachers in senior secondary colleges carry out the moderation with Subject Group Leaders support in a guidance capacity. All senior secondary teachers are involved in this process. These judgements are backed by statistical moderation which compares scores from student assessment placed on the same scale as results from the ACT Scaling Test.

Anecdotal evidence suggests that moderation does occur at a school level, however the absence of a system-wide curriculum framework means that different schools do different things. There is some evidence that when schools work closely together in cluster groups some discussion and moderation activity does occur between schools, but this is not widespread, except for Years 11-12 as noted above.

Evidence of Improvement

The practice of school-based curriculum development means that currently approaches to consistency of teacher judgements is a school-based, or in some circumstances, a cluster-based priority. No data is collected as to the efficacy of any consistency measures that might be in place in these contexts.

The senior secondary moderation model is well established. There is no evidence of improvement in consistency of teacher judgements per se, after each moderation day reporting of the consensus view of the presented work is fed back to teachers and schools and is used as the basis for identifying system curriculum and assessment issues by the ACT Board of Senior Secondary Studies. This feedback through the system is one of the strengths of the ACT senior secondary approach.

* * * * *

NEW SOUTH WALES

Context

The K-12 Curriculum in NSW is overseen by the NSW Board of Studies (BoS). Standards are made explicit in the Board syllabuses. The curriculum is organised in stages. Early Stage 1 to Stage 3 covers the years from Kindergarten to Year 6. Stage 4 covers Years 7-10 and Stage 5 covers Years 11 and 12, the Higher School Certificate. For Primary schools in NSW there are 24 Foundation Statements which cover the six primary syllabus areas and four Stages from K-6. Foundation Statements describe the state-wide common curriculum requirements and prioritise what needs to be taught in all primary schools. They set out a clear picture of the knowledge, skills and understanding that each student should develop at each stage of primary school.

The focus of strategies for improving consistent judgement in NSW is on developing teacher knowledge and understanding of the standards against which they are assessing their students, collaborative planning with clearly stated learning objectives, and on validation of teacher judgements through social moderation. Generally the approaches used have involved providing teachers with professional learning about curriculum planning and social moderation supported by online or printed materials including work samples and sample assessment activities. This has been done more extensively in the K-6 area and is now tied to A-E reporting requirements. These steps address the key elements of building knowledge and understanding of the standards, assessing student work against those standards and validation of teacher judgement extensively. Support for teachers in building knowledge of the curriculum standards is available and linked with planning. Evidence of student achievement is provided in the form of exemplar assessment tasks, and annotated student work samples assist teachers to recognise the kinds of evidence of students' work that demonstrate achievement.

Approaches to Consistency

Supporting knowledge of the standards

The Curriculum Support Directorate provides advice and support to teachers and schools regarding aspects of the K-10 syllabus and provides workshops and other professional learning for teachers to enhance their knowledge and understanding of current syllabuses.

The K-6 Foundation Statements were designed to assist teachers identify and prioritise what is taught in their classrooms and highlight at a broad level the key standards and outcomes of the syllabus. These clear, broad statements of what students are expected to learn and achieve has

helped teachers in the primary area build their knowledge of the NSW primary syllabus and drawn attention to key outcomes particularly effectively.

Standards are made explicit in the Board syllabuses as standards articulating aims and objectives, outcomes, content, and Stage/Foundation statements. These syllabus standards focus on what is learned, and performance standards focus on how well students have learned. The performance standards articulate a common grade scale, assessment activities and annotated student work samples.

Curriculum planning

Commencing in 2006 and continuing through to 2008 a number of professional learning opportunities focussed on planning, programming, assessing and reporting processes have been and will be offered to teachers. Workshops on joint planning for student learning, collaborative planning of assessment, moderation within and between schools and collecting and sharing evidence of student achievement are intended to support the implementation of the Curriculum Planning and Programming, Assessing and Reporting to Parents K-12 policy.

Planning, particularly collaborative planning, is highlighted in the materials developed by the Department of Education and Training as a key to developing consistency in what is intended to be taught and developing common understandings of assessment practices and tasks which accurately reflect the syllabus. PowerPoint presentations for each subject area advise teachers to share their teaching program with colleagues in other subject areas and collaboratively develop understandings of assessment practices which are based on syllabuses.

Evidence of achievement

In 2006, no central bank or collection of common assessment tasks appeared to be available to assist teachers in helping to build the body of evidence against which they make judgements. Teachers are encouraged to plan assessment processes and tasks with colleagues, which is practical on a small scale but less so on a larger state or system-wide approach.

Assessment of evidence against the standards

To assist teachers to assess students against the standards, a number of work samples are available through the Assessment Resource Centre website. Each work sample is accompanied by a grade commentary and description as well as the Foundation Statement in the K-6 area and an area for assessment link for Years 7-10. The activities and work samples on the website are sourced from Board of Studies publications, teachers who participated in grade alignment workshops and others who were invited to contribute. All samples are authentic exemplars of students' work.

The samples have been aligned to the A-E reporting scale and organised in such a way as to reflect achievement in each stage on that scale. They were subjected to a rigorous grade alignment process by the Office of the Board of Studies using the resources of the Department of Education and Training, the Catholic Education Commission and the Association of Independent Schools who nominated teachers from their sector to carry out the process of aligning each sample to a grade level. The work samples on this website can be considered benchmark exemplars of student work which meets the syllabus requirements and grade allocation criteria.

Validation of teachers' judgements

New consistent teacher judgement programs are informed by earlier NSW QTP and AGQTP projects Consistency of Outcomes and Stages Assessment in Literacy and Numeracy (COSA) and Consistent Teacher Judgement Online. The goal of COSA was to support the implementation of materials to assist teachers to assess student achievement in literacy and numeracy against stage

outcomes and standards from Foundation to Stage 3. Schools were supported for within-school and across-school professional dialogue by district teams which promoted sharing across and between districts. Online materials including examples of best practice were also made available. The focus of this program was facilitating social moderation amongst teachers across the state system.

Consistent Teacher Judgement Online, which followed the COSA project, focussed on developing consistent judgement among teachers in rural and remote regions. Participants in the program used WebBoard to communicate with colleagues in rural and remote locations across NSW. They compared professional judgements of students' work in order to confirm or adjust them with reference to syllabus outcomes for English K-6. The goal was to facilitate social moderation opportunities but teachers were also provided with work samples in English and mathematics and some additional resources for science and technology. These materials were designed to support teachers in building and selecting from their students' work a sufficient body of evidence on which to make their assessments. Collaborative planning was also built into the online course and included the development of common assessment tasks. Implicit in the program was the notion that knowledge and understanding of the syllabus was essential to the moderation process, but that in turn moderation promoted improved understanding of the syllabus.

Targeted Resources and Professional Learning

The Consistent Teacher Judgement in Action resource was developed in 2002 to support teachers to ensure that consistent teacher judgements about student achievement can occur across all key learning areas. In 2006 as part of quality assurance procedures, the original consistency, moderation, assessment and reporting materials were reviewed and updated for website publication. The new Consistency of Teacher Judgement website went live in mid-October 2006. (<http://www.curriculumsupport.education.nsw.gov.au/consistent-teacher/index.htm>) and will be maintained and updated regularly. This is a relatively cost-effective means of providing teachers and schools with assessment and consistency advice and access to professional learning tools.

Four key questions were used in developing the website: What do you want the students to learn? Why does that learning matter? What are you going to get the students to do (or produce)? How well do you expect them to do it? Materials to support the 'Consistency in practice' and 'moderation in practice' workshops which are intended to be run within schools are collected on the website. Six PowerPoint presentations, developed for specific KLAs focussed on using the syllabus for consistent judgement are available to schools and are intended to be used at staff meetings, stage meetings or staff development days. The 'moderation in practice' workshops are resourced with sample planning grids, a matrix for the evaluation of current school practice in addition to the PowerPoint documents. Work samples for the Connected Outcomes Groups (COGS) will be added to the website in future.

Many practising teachers are engaged as markers for the annual Basic Skills Test, and the English Language and Literacy Assessment and this creates significant professional learning opportunities, extending knowledge of assessing student writing.

Independent schools in NSW have also focused on improving consistency of teacher judgement. Professional learning workshops have been offered by AISNSW and these workshops have seeded school based professional learning sessions. Teachers are offered training in assessment for learning and tracking performance to outcomes and are then asked to bring work samples for moderation and the development of a marking system which corresponds to criteria or rubrics depending on the school context. Pilot workshops have been conducted with three independent schools, each school at a different point. Further workshops are being offered in 2007.

Evidence of Improvement

There is little explicit evidence that the approaches outlined above and used in NSW have improved consistency of teacher judgement. The approaches enhance the capacity of teachers to work together to build knowledge of the syllabuses, assessment practices and tasks and moderation but the results of such activities have not been tracked in any coherent manner. Anecdotal evidence and the continued development of consistency programs in NSW implies that over the course of time the various programs in place in NSW have had an impact on improving consistency of teacher judgements.

Summary

The Consistency of Teacher Judgement website is a comprehensive, accessible and continually updated source of information for teachers seeking to develop a better approach to consistency in their assessment practices. Drawing on the work previously done in NSW to build greater consistency, it provides resources in a practical and accessible medium. Collecting all resources and information which related to consistent teacher judgement in the one place highlights the role that each element plays in maintaining consistency more effectively than resources available from a range of different agencies or sources. However, like all web based resources it is difficult to determine the extent to which it is being utilised in schools and by teachers other than through anecdotal evidence.

The NSW approach to moderation is localised, focussing mostly on what schools can do to achieve consistency within the school context. The syllabuses in NSW are state-wide, but the approach to moderation is on a smaller scale. There is no central scrutiny of what process is used for these localised moderation activities so it is difficult to determine how well the advice offered by DET is implemented.

The NSW Board of Studies Assessment Resource Bank is a particularly effective tool for teachers to build consistency through using annotated, benchmarked student exemplars to inform their own assessment practices. The Assessment Resource Bank provides detailed analysis of student work samples against the syllabuses in a format that is simple to follow. The depth and breadth of this resource is another of its strengths.

* * * * *

NORTHERN TERRITORY

Context

The Northern Territory Curriculum Framework (NTCF) is the foundation curriculum document for schools in the NT. The NTCF identifies learning outcomes for students from Transition to Year 10 in eight nationally agreed learning areas in addition to the EsseNTial Learnings, learning technology, ESL and Indigenous languages and culture. The curriculum framework is arranged into Key Growth Points at school entry and Bands, each of which cover two years of schooling. It is designed to support schools to develop flexible teaching and learning programs that are inclusive of varied pedagogical approaches, assess learner progress and report on the outcomes achieved. The framework is supported by series of system-developed materials, system-recommended resources and teacher recommended and devised resources through the NT Schools Portal.

The NT approaches to consistency focus on validation of teacher judgements, assessment task development and assessment against the standards. There are plans for future work supporting teachers in their planning arrangements. While there is a culture of moderation in the NT there are no formal mechanisms for measuring any improvement in consistency of teacher judgement.

Extensive work has been done to familiarise teachers with the Curriculum Framework and to build knowledge and understanding of how to plan opportunities for students to demonstrate achievement of those standards.

Approaches to Consistency

Supporting knowledge of the standards

The NTCF is well supported through the NT School Portal, a secure web site that offers a broad array of resources and services, such as e-mail, forums, access to shared resources, and search engines. Knowledge and understanding of the standards, as in most other jurisdictions, is reported to be strong.

Evidence of achievement

The Assessment for Learning project which is funded from the Indigenous Education Strategic Initiatives Program contains an outcome addressing quality of teacher judgements and improved quality assessment strategies, design and selection. This assists in the design of a body of evidence from which to make consistent judgements about student achievement. There is little published detail available about this project and how the outcome is addressed in practice.

Assessment of evidence against the standards

Annotated samples of student work are available from the Department of Employment, Education and Training and have been used in school and cluster moderation; however schools often provide their own exemplars. This is the preferred method in the NT as it connects more closely with the work being done in schools. There is no centrally available bank of annotated student work samples as are available in some other jurisdictions.

The introduction of A-E reporting has sharpened the focus on consistent judgement in the NT. The Curriculum eTool (CeTool) software operates at school level to monitor student achievement. Schools enter student achievement information for individual students. The CeTool is now linked to A-E reporting and generates an A-E report with the goal that teachers remain focussed on the outcomes in the Curriculum Framework rather than on attempting to assign a grade to student work. Schools are expected to, and do, moderate teacher judgements before entering data into CeTool.

Incorporated into the Building Better Schools program is the Standards Validation Project. A model for system-wide standards validation is being developed for implementation in 2007 and involves school-based moderation, across cluster moderation and NT Standards Validation. Standards validation and assessment workshops designed to give NT teachers guidance and support in making rigorous and comparable assessments of student have been offered. They provide teachers with assessment tools and opportunities to take part in validation procedures. A trial validation process was undertaken with English teachers with the intention of developing procedures for system validation processes in English and mathematics. At assessment workshops teachers are supported to design assessment tasks for students across the range of Bands in the NTCF and to develop school based moderation processes. English and mathematics teachers who have participated in system-level validation workshops have worked on verification of teacher judgements. Samples of student work are presented at these workshops and annotated and published as exemplars of the standards at each band level.

Validation of teachers' judgements

There has been a culture and expectation of moderation in schools in the NT since the introduction of the National Statements. Moderation sessions and practices in place at this stage are mostly

school based, however some clusters have offered activities. For example, one cluster uses a model which is being promoted by the Department where representatives from schools attend cluster moderation sessions and take the learning from the session back to their schools. The absence of a formal moderation system in the NT makes it difficult to assess the efficacy of the moderation activities which occur there.

Implementation of a moderation regime faces particular hurdles in the NT, for example, supporting the large number of remote schools. Schools in remote areas with just one or two teachers are grouped geographically with a group principal who will involve teachers from the group in moderation activities. However this is not always practical. The push to move resources and moderation online attempts to redress the difficulty of bringing remote teachers into the centres for this kind of work. Moving resources and activities to the online environment makes them more accessible to remote teachers, however it is difficult to gauge whether the resources are used as intended or in a systematic manner.

Anecdotal evidence reported to the Department suggests that moderation and consistency has been better implemented from beginning of school to Band 3 level, the primary school years. It has been observed during cluster moderation and anecdotally that evidence of learning is conceptualised differently by primary and secondary teachers. For example, primary teachers consider draft work whereas some secondary teachers tend to look only at published or completed student work. In addition, there is a large jump in outcomes between Band 3 and Band 4 which mirrors change and movement of students from primary to secondary school. This gap has been identified and is being addressed in review and helped by the introduction of middle schools which move Year 7 students from a primary context to a 7-9 environment. Greater support for secondary teachers is often required as teachers from different subject areas may find fewer colleagues with whom to work when attempting to participate in moderation activities.

The Multilevel Assessment Program (MAP) assesses reading, writing, spelling and numeracy levels for students in Years 3, 5 and 7 in the NT. MAP writing tests are marked and moderated by a team of current teachers who are released from schools for the marking period. Marking is done three times and moderated on these occasions. These teachers gain valuable experience moderating student writing and take enhanced knowledge and experience back to their schools.

Moderation protocols are currently under development and are influenced by the protocols being used in other states. The Heart of Teaching protocols have also been drawn upon during the development process. The final protocol is likely to be a hybrid version of what already exists in other jurisdictions, contextualised for the NT.

A range of online moderation and assessment resources is planned, particularly to assist teachers and schools in rural and remote areas in the Territory.

Targeted resources and professional learning

Professional learning that incorporates the key elements of consistent judgement are often embedded in projects being undertaken by the Department – for example, the Literacy Project and Early Years Framework. At present there are no dedicated consistency or validation resources or professional learning available to teachers in the NT, however the plans outlined above provide scope for such targeted work.

Evidence of Improvement

The processes in the NT are still largely under development. There is no clear evidence of improved consistency from the approaches detailed above, however anecdotal evidence of improved moderation practices is plentiful, particularly in those clusters where moderation and consistency has been a focus.

Summary

Where school clusters have highlighted it as a priority, consistency through moderation is undertaken regularly and draws participation from teachers in schools who participate in cluster moderation and then return to their schools with new knowledge to be shared with other teachers. Moderation and validation of teacher judgement is not a priority in every cluster however, so implementation territory-wide has not occurred.

The way in which the CeTool has been implemented has helped establish the importance of assessment referenced clearly to standards in the NT Curriculum Framework. The NT DEET has effectively incorporated the A-E reporting regime into this tool yet maintained the focus on the NTCF as the base from which assessments are made, rather than the A-E scale.

* * * * *

QUEENSLAND

Context

The *Queensland Curriculum, Assessment and Reporting* (QCAR) Framework will define what Queensland school students should learn, how they are assessed, and how schools report student achievement. The Framework will define essential learnings, standards, assessment strategies and reporting requirements in Years 1–10. The initiative will be implemented in all Queensland schools by 2008. The timeline for development is shown in Table 2.

Table 2: Timeline for QCAR development

	2005	2006	2007		2008
Essentials and Standards	Identify and describe Essential Learnings and specify Standards	Develop and refine Essential Learnings and Standards with identified partner and trial schools	Provide draft of Essential Learnings and Standards for open trial and implementation	Publish and disseminate final version of Essential Learnings and Standards	Full suite of P-12 curriculum framework is available, including Essential Learnings, Standards and assessment resources and tools
Assessment	Develop, collect and refine items and resources for Assessment Bank	Trial statewide assessment model in English and Maths for Years 4 and 9	Work with selected schools and teachers to develop and refine further tools and resources	Develop and refine quality assurance processes for assessment tasks	
Reporting framework	Develop, trial and refine reporting framework			Publish and disseminate reporting framework	

Source: http://www.qsa.qld.edu.au/qcar/index.html#what_is Accessed 20 February 2007

The policy direction for the framework was developed by the Department of Education and the Arts in collaboration with the Queensland Catholic Education Commission, the Association of Independent Schools Queensland, and the Queensland Studies Authority. The Framework defines the essential curriculum for all students P – 10 and sets standards of student achievement. Assessment uses a five point scale, A-E, and the standards will be used by teachers in the state-

wide assessment of essential learnings in Years 4, 6 and 9. A key goal in implementing the new curriculum framework will be to achieve alignment between curriculum, assessment and reporting.

Until the Queensland curriculum is fully implemented, moderation processes including key approaches to consistency such as building knowledge of the curriculum and standards and assessment against standards will be limited. The state does however have a strong background in moderation at the senior secondary level, and, from the New Basics Trial, considerable experience of moderation in the compulsory years. This experience will inform the development of approaches to developing consistency approaches for teachers in the primary and junior secondary areas.

Approaches to Consistency

Supporting knowledge of the standards

With the implementation of the new framework, teachers will be required to match evidence of students' performances to the standard descriptors. To assist this process annotated illustrations of student achievement for the standards will be available. School authorities will have responsibility for implementation of the framework.

Curriculum planning

The Queensland Studies Authority will develop common assessment tasks at Years 4, 6 and 9 to assess student achievement against the standards in English, Maths, Science and one other area. Data from these tasks will be reported at the system and individual student level. The bank of rich tasks produced for the New Basics project can continue to be used with the QCAR framework and new blueprints for designing rich tasks are assisting teachers to build their own task design skills.

Evidence of achievement

The New Basics Trial in Queensland placed significant emphasis on building a body of evidence from which teacher judgements about achievement were drawn. The nature of the rich tasks used for assessment meant that assessing from a body of evidence was built into the task regime. Rich tasks detail the 'assessable products' that students create in the course of meeting the completion. Schools that have participated in New Basics are in a good position to continue to build a strong body of evidence prior to assessing student achievement against the Standards soon to be introduced in Queensland, and this has been recognised by QCAR.

Assessment of evidence against the standards

In Queensland a new state-wide assessment regime will be introduced. Teachers will judge these assessment tasks using marking guides provided to them which illustrate the distinguishing features of student achievement. It is intended at three key points in the P-10 years that comparable assessment against the standards will be done with the intention of ensuring consistent teacher judgements of student achievement. This process requires the use of common assessment tasks.

An assessment bank is being developed for teachers. The assessment bank will contain high quality assessment tools for P-10 that are linked to the essential learnings and standards. Teachers will be able to use the resource as they choose to complement their own assessment practices and strategies. Every assessment tool will be accompanied by an administration and marking guide, links to teaching and learning, and student work samples illustrating what is needed at each level described in the marking guide.

Validation of teachers' judgements

Moderation of teacher judgements has a long history in Queensland in the senior secondary years. The assessment system of moderated school-based assessment in the New Basics Trial provided a strong model for the compulsory years of schooling through the moderation of the rich tasks. Currently no other moderation approaches are in place, although there are plans for the future.

In 2002 the Assessment and Reporting Taskforce concluded that design of assessment tasks and moderation as a professional development device were two areas which some teachers and schools needed to attend to. The report recognised that moderation not only addresses teacher judgement but leads to conversations about teaching and learning and curriculum knowledge. The need to develop of electronic means of moderation was highlighted for schools that are small and/or isolated.

Social moderation has been the preferred model in Queensland, rather than statistical moderation and moderation by visitation. The New Basics moderation model consists of a 'guild of professionals', that is teachers, their peers in the same school and in other schools, and teams of moderators who act as facilitators for the moderation sessions themselves. The moderation process begins with a meeting to establish a common understanding of the standards against which the task is being assessed and making a provisional grade allocation based on discussion and close examination of the work and any secondary evidence which might assist. The process then moves on to ratification of that grade assignment. Ratification involves collecting together samples of student work which represent the range of grades awarded by the school including the highest grade and the lowest grade. Task sheets and secondary material used in completion of the task accompany these range samples, as do the annotated grading master for each task in the collection. Moderators then review the work and either ratify the provisional grade or return the tasks for regrading if state-wide comparability cannot be achieved should the grade stand.

Once confirmation of grades is achieved the school submits the total distribution of grades to the ANB Branch as well as all ungraded results where the task has not been completed. At all stages of moderation both the teachers in the school, their peers and the moderation team are involved. Teachers involved in the New Basics trial described the moderation process as one of the most important contributors to professional skills enhancement and to developing confidence in applying the model for grading students' Rich Task performances (Matters, 2005).

Targeted resources and professional learning

The Catholic sector in Queensland has also implemented strategies directed towards the development of a system of consistent teacher judgement in schools. Schools working in clusters supported principals and teachers to address curriculum change in 2001/2002. The goal was to develop consistent teacher judgement in relation to the outcomes of the Queensland School Curriculum Council Syllabuses and the Learning Framework for Catholic Education in the Archdiocese of Brisbane. In 2001 Brisbane Catholic Education produced the *2001 Consistency of Teacher Judgement Information Kit* and schools used that resource in conjunction with the Consistency of Teacher Judgement CD-ROM to develop a process of moderation and professional learning for Brisbane Catholic schools. A program of *Consistency of Teacher Judgement* has been in place in the Brisbane Archdiocese since 1995 and supports the development of common understanding of curriculum and allows for inter-staff and intra-staff meetings for social moderation.

Lutheran Education Queensland has also investigated consistency of teacher judgement and developed policy acknowledging its importance. The policy identifies four broad strategies for developing consistency: shared understanding of the syllabus framework and outcomes; collaborative planning; comparable assessments; and agreed demonstrations of learning.

Evidence of Improvement

Evidence of improvement in the quality of teacher judgements can be found in the New Basics Research Report. The report of the trial of the New Basics assessment system indicated that it did produce valid and reliable assessments, and a reasonable degree of comparability in the grades awarded and reported on (New Basics Research Report, 2004). The report noted that no elements of the New Basics assessment system could be dropped without a decrease in effectiveness. These five elements were identified by Allen (2003) and are discussed in more detail in Chapter 3 of this report.

Summary

The structured moderation process developed for the New Basics project provides a model for moderation which incorporates the professional learning aspects valued by teachers when they are able to discuss student work with their peers with a level of oversight provided by the moderation teams and submission of grades for ratification and eventual reporting to the central body. The moderation teams help to ensure that teacher judgements and grades allocated to student work were consistent across the schools participating in the New Basics Trial. This combination of intra-school, inter-school and specialist moderation provided a high quality checking system.

* * * * *

SOUTH AUSTRALIA

Context

The *South Australian Curriculum Standards and Accountability Framework* (SACSA) was introduced in 2001 and outlines the learning outcomes from birth to Year 12. There are four bands, the Early Childhood band from birth to Year 2, the Primary Band for students in Years 3-5, the Middle Band, Years 6-9 and the Senior Band, Years 10-12.

Within SACSA, Standards represent expectations of learners, and provide a common reference point for educators to use in monitoring, judging and reporting on learners' achievement over time. In the Early Childhood band, Curriculum Standard 1 is not used until children begin school in the Reception year; before children start school planning for learning is based on the Developmental Learning Outcomes. The Curriculum Standards 1-5 cover the Years R-10, and Year 12 Standards for senior secondary students include externally accredited curriculum. At the time of this investigation, in late 2006, there was no required reporting against state-wide standards. For 2007, the expectation is that all schools will formally record SACSA data in at least one learning area. Schools have been encouraged to record data for English, Mathematics or Science. New web-based software, the SACSA Achievement System (SAS), is designed to allow schools to analyse the data they enter.

Moving Forward with SACSA is a strategy developed in 2004 to refocus and to support all schools and pre-schools in the effective use of the SACSA as the core curriculum policy of the DECS. A key objective was to assist all sites to continue to improve the quality of the engagement and outcomes of learners, and to support schools' improvement in SACSA Standards. The strategy continues in 2007.

Within the South Australian Department of Education and Children's Services (DECS) a manager and small teams of policy officers, the SACSA Team, and the Standards Team, developed resource material and provided support to districts and schools. The work was targeted, and included a specific focus on consistency of professional judgement activities, providing professional development on consistency of teacher judgement. A CD-ROM provided a range of resources to

support the understanding and practice of consistency of teacher judgement. It includes a moderation PowerPoint and the process that has successfully been used with schools from R-12.

The *Moving Forward with SACSA* approach to consistency focuses on four key concepts – a common language for describing student achievement; recognition that students demonstrate knowledge, skills and understandings in diverse ways across a range of contexts over time; collaborative moderation of ‘sets of evidence’ to confirm teacher judgements; and using SACSA Standards as the common reference point against which judgements about student achievement are aligned.

Approaches to Consistency

Supporting knowledge of the standards

The SACSA website is a detailed resource which supports teacher knowledge and understanding of the curriculum, standards and accountability regime in South Australia. Professional learning modules, also accessible online through the website, provide further support on particular aspects of the framework. In addition to the online offerings, there are strong district-based measures to support the framework. There is a central SACSA contact person provided to assist with district planning, district-based professional support services and special consistency and data activities organised in each district.

One of the principles articulated in the South Australian *Consistency of Teacher Judgement* professional learning module is a common language for discussing, interpreting and understanding outcomes in the framework and this can only be achieved when teachers know the curriculum and have opportunities to talk together about it. The professional development workshops include a strategy whereby teachers analyse the language of the standards in detail, and identify the key knowledge and skills

Curriculum planning

Collaborative planning is a key element in the SACSA advice provided to teachers in professional development materials and recommended in the consistency of teacher judgement materials as an important measure for promoting consistency.

Evidence of achievement

It is implied in all SACSA consistent teacher judgement material that evidence of student achievement is to be drawn from a range of contexts over time.

Feedback on consistency workshops indicated that teachers encountered different ways of teaching, and developed insights into the need to look closely at the outcomes, to link assessment tasks to outcomes, and to gather a broader range of evidence than writing.

Assessment of evidence against the standards

Assessing SACSA Outcomes Through Standards (ASOTE) student work exemplars are provided on the *Moving forward with SACSA: Resources for developing consistency in teacher judgement* CD-ROM. Each exemplar is accompanied by a plan for teaching and learning, the student evidence itself, commentary and reflective questions to assist teachers assess their students’ work against the SACSA standards. The evidence is presented as sets with three or four samples of work for each task. This is a clear indication of how work completed over an extended period of time in a range of contexts can be used to build the body of evidence upon which the final judgement of achievement is made. Each set of ASOTE student work exemplars includes a statement addressing which standards have not been achieved in each set.

Validation of teachers' judgements

Moderation of evidence to confirm judgements is encouraged and builds on the work done in the earlier project. Collaborative moderation and triangulation has been done by individual teachers against their own records of student achievement in the past. Collaborative moderation follows the SACSA moderation protocol. The protocol is derived from the 'Tuning Protocol' developed for the Coalition of Essential Schools. The protocols are seen to be significant in the moderation process, especially for the first time teachers meet to moderate.

School leaders are offered professional learning sessions to enable them to conduct their own school and regional moderation sessions. 'Developing consistency of teacher judgement' is a three-session professional learning module, the second in a series of Assessment and Accountability as Learning Opportunities modules provided by officers from DECS. It provides an opportunity for teachers to explore the principles of consistent judgement and reflect critically on the process of making judgements while participating in activities which support consistent judgement.

This professional learning module emphasises the development of a common language between teachers for discussing and assessing student work, diverse demonstrations of achievement, social moderation, a range of evidence from an extended period of time and using the Standards as the common point of reference. The PowerPoint document used to deliver this professional learning module is accessible online and school leaders are encouraged to use the resources from the website to develop professional learning for staff which is appropriate to their circumstances and contexts. In the consultation, it was noted that the demand for workshops had increased following the introduction of the A-E reporting requirements.

Targeted resources and professional learning

A number of resources have been developed to provide support for teachers in using the SSACSA Standards, and to build capacity in making consistent judgements. These include the ASOTE exemplars, the newsletter, a moderation video (includes protocols), PowerPoint for leaders, SACSA CD11, and advice on unpacking the standards.

The Consistency of Teacher Judgement CD-ROM³, was made available to all schools in South Australia. It is accessible by individual teachers and schools; however, there is little information to show the extent to which it is used in schools.

The *Moving forward with SACSA: Resources for developing consistency in teacher judgement* CD-ROM is also available to teachers in all South Australian schools. This is a more recent publication, linked more closely to the South Australian context. Uptake of the resource is voluntary and requests to the Department for further support have been increasing indicating that the materials are being accessed and used in schools.

Currently, the Department of Education and Children's Services produces a newsletter, *Let's Talk Assessment*, the first issue of which looked at consistent judgement and moderation. In this newsletter, professional judgement was succinctly described: *Professional judgement is an estimation by a teacher of a student's achievement of SACSA Outcomes and Standards, using evidence of learning collected over time and in a range of contexts*. Elsewhere in this first issue, attention is drawn to the need for a common language³ for describing student achievement as

³Produced by the South Australian Department of Education, Training and Employment in collaboration with the Victorian Department of Education, Employment and Training and the Queensland School Curriculum Council, under the Commonwealth *Quality Outcomes Programme*, in 2000.

fundamental to consistency of teacher judgement. The newsletter is intended to keep assessment issues and practices at the forefront of teachers' minds.

Evidence of improvement

Feedback from moderation meetings indicated that teachers were aware of their lack of prior experience in examining students' work against standards, and had gained new understandings from the workshops.

Summary

The *Moving Forward with SACSA* strategy has focused on building consistency of teacher judgement in South Australia and has developed a key collection of resources, materials and professional learning modules which are easily accessible to teachers in the State. Items such as the *Let's Talk Assessment* newsletter keep issues of consistency and assessment high on school and teacher radars.

While the program for developing consistency of teacher judgement is largely voluntarily implemented in schools, anecdotal evidence from the central office indicates that more schools are addressing issues of consistency and requesting further advice and assistance from the central team.

The support for moderation processes, including the South Australian moderation protocol means that social moderation between teachers in schools is quite strong. In places where the materials have had greatest impact there is anecdotal evidence that moderation processes have become incorporated as a matter of course in the teaching-learning cycle, and the moderation protocol providing a strongly structured basis which schools use to meet their own particular needs.

* * * * *

TASMANIA

Context

The Essential Learnings Framework provides the curriculum for learning in Tasmania from birth to 16. There are five organisers in the Essential Learnings Framework: thinking; communicating; personal futures; social responsibility; and world futures, which comprise 18 Key Elements. There is a range of outcomes for each key elements ranging from Standard 1 to Standard 5, encompassing the period children are in kindergarten to Year 10. Tasmanian schools were required to start reporting and assessing against the Essential Learnings standards from 2005.

The Quality Moderation of Assessment Process (QMAP) was initiated in 2005 to provide assistance to all teachers in K-10 classrooms to make consistent judgements about students' performance against the outcomes and standards of the Essential Learnings Framework. QMAP operated throughout 2005-2006 across Tasmanian government schools. Resources available for QMAP included funding for three Principal Education review officer positions, funding for 30 Assessment Support Leader positions, pupil free days, the Guiding Assessment Tasks, collections of work samples, and one day's relief teaching per year for all schools. In summary, the QMAP resourcing included staff, teacher time, and materials.

Approaches to Consistency

Supporting knowledge of the standards

The Tasmanian approach involved an extensive program of moderation and validation of teacher judgements and is acknowledged to have generated a depth of knowledge and understanding of the curriculum framework in Tasmania that would not have existed otherwise.

The *Learning, Teaching and Assessment Guide* (LTAG) is a developing resource that supports educators implementing the *Essential Learnings Framework*. The LTAG was developed from the work of teachers, schools and curriculum personnel as they explored the new curriculum framework and considered how practice might change as curriculum implementation proceeded. An *Essential Learnings Assessing Guide* was published in 2005, and materials to support this guide are available online on the LTAG website:

<http://ww.ltag.education.tas.gov.au/assessment/default.htm>

Evidence of achievement

QMAP incorporated the development of guiding assessment tasks. These were developed centrally to assist in building consistent teacher judgement of standards and outcomes at a state-wide level. Guiding assessment tasks are common tasks developed collaboratively by teachers, School Education Division and Office for Educational Review staff. They address the calibrated key elements and are administered to entire cohorts of students. The assessments are marked externally and reported centrally to be incorporated into the SARIS database. Results are reported back to schools and schools are encouraged to use the tasks in their own moderation processes and to compare the assessments made internally with those made centrally. These assessments provide a system-wide quality control mechanism and feedback for teachers to help build the body of evidence against which they determine achievement of the standards for students across the whole range of their work.

Assessment of evidence against the standards

In addition to the whole cohort guiding assessment tasks, sample tasks have been developed and are used in schools on a voluntary basis. Students complete the tasks which are submitted along with explanations for assessment decisions in relation to the sample for external review. Practising teachers are involved in the central assessment process. The results are reported back to schools. The intention of the sample program is to build the capacity of teachers to apply the curriculum standards in their assessment processes consistently.

Validation of teachers' judgements

QMAP is designed to build teachers' capacity to make consistent judgements about student work that are valid, reliable and fair and to provide parents with an assessment of their child's progress against calibrated standards. It combines three main components – cluster based professional learning, school-level moderation, and guiding assessment tasks.

Significant resources were available to QMAP in 2005-2006, including a support team in the Office of School Review, Assessment Support Leaders appointed to local school clusters, and time for state-wide, cluster and school moderation days. Three student free professional learning days were allocated for QMAP, two of the days facilitated by the department and one day allocated for schools to use for internal school level moderation. The days facilitated by the Department were run through school clusters and focused on collaboratively assessing student work using an assessment protocol.

The three moderation days were used differently in 2005 and 2006. In 2005, the first day was a cluster day, the second day a school-based day, largely focused on work samples, and the third day was a cluster day. In 2006, the first two days were cluster days, and the third day a school-based days. Support will be offered again in a scaled back form of QMAP in 2007, at both a cluster and school level.

The *Tasmanian Essential Learnings Assessment Protocol* was developed, with reference to the *Tuning Protocol* developed for the Coalition of Essential Schools, to provide structure for teacher discussions around student work that occur during moderation. It was designed to enable teachers to make judgements in a manageable and consistent way against the standards and is used to ensure that the focus for moderation is determining an outcome through careful consideration of the available evidence.

The moderation days involved developing understanding of standards, progression and support material as well as looking collaboratively at student work and compiling evidence sheets describing achievement against the standards. Planning for common assessment tasks was also incorporated. In 2006 the first moderation day was at the cluster level and focussed on the middle school Years 5–8, allowing teachers from different schools to collaborate. In cases where there was only one secondary school in a cluster opportunities to collaborate with schools in neighbouring clusters were provided. Further moderation days were held at a school level to provide an opportunity for teachers to develop consistency of judgement in relation to outcomes and standards being reported by the school and to assist in embedding moderation processes in each school's organisation.

One of the findings from the 2005 implementation of QMAP was that most benefit was derived from these exercises when teachers provided the samples of work used by the groups. As new key elements were introduced into the moderation process multiple work samples were provided to teachers to help build common understandings of the standards. Teachers were encouraged to use a broad range of evidence when making judgements and were encouraged to also bring non-written samples to moderation.

Moderation work was supported by 30 Assessment Support Leaders working in clusters. The Assessment Support Leaders had some training in coaching and attended a workshop with Lorna Earl from Ontario along with principals and personnel from the Office of School Review in 2005. In addition to this cluster level support, QMAP was backed by three Principal Education Review Officers from the Office of School Review in a central, state-wide capacity.

Evidence of Improvement

Evidence of improvement in consistency of teacher judgements is primarily anecdotal. QMAP personnel in the Office of School Review noted that teacher discussion around student work had changed, and assessment practices had changed. The modelling of teacher judgements against school data from state-wide testing had provided a guide to levels of state-wide consistency, and to patterns within schools. Where there were discrepancies identified in the modelling, Principal Leaders could discuss these with the principal of the school. The modelled data helped to contribute to the consistency of teacher judgements, helping teachers to see whether they were 'in the ball park'.

Initially, in moderation workshops, teachers' assessment of the same set of work samples across the state demonstrated very little consistency, but there appeared to be improvement over time. Teachers also became more willing to bring samples of their students' work to moderation meetings. Feedback from teachers indicates recognition of the need for moderation, and schools have introduced moderation processes. Many teachers were prepared to facilitate moderation

sessions, and to provide work samples. Feedback on moderation sessions was collected by the Department, and the following observations and comments were made:

- Multiple work samples from one student were used
- All teachers brought work samples, thus providing authentic samples from members of the group
- Teachers were more relaxed about the moderation process and able to use a variety of formats for discussing student work
- Meeting with colleagues to look at assessment from across a range of grades including primary and secondary was a positive experience
- There was consensus between school and cluster moderation.

Summary

The coordinated activities and support provided through QMAP addressed issues of consistent teacher judgements in Tasmania. The coordinated approach and specific allocation of resources provided extensive support over two years to all schools.

* * * * *

WESTERN AUSTRALIA

Context

The Western Australian Curriculum Framework provides outcomes and standards across the years of schooling. The *Outcomes and Standards Framework* provides developmental sequences for the Learning Area Outcomes in the Curriculum Framework. These are continua of increasing student understanding and skill development across eight levels. The levels are not linked to age or school year levels. In addition, achievement targets exist at Years 3, 5, 7 and 9. Achievement targets are set for outcomes deemed critical for students in those year levels. In WA standards are assessed through the judgements of teachers, along with the state-wide testing program.

Professional learning about the collection and analysis of quality literacy assessment information is prominent in many recent Department of Education initiatives, including *Getting it Right, First Steps, Making Consistent Judgements, Literacy Net, Aboriginal Literacy Strategy* and *Literacy and Numeracy in the Middle Years*.

A considerable investment has been made in Western Australia towards ensuring consistent teacher judgement across the state. The *Making Consistent Judgements* professional learning activities in 2004-2005 concentrated on the key elements of teacher understanding and knowledge of the standards, assessment against the standards and validation of judgements through social moderation. Particular attention was paid to planning and selecting student work through workshop time devoted to collaborative planning for the collection of evidence. Planning for opportunities for students to demonstrate achievement of standards is not heavily emphasised but is implied through the planning for collection of evidence.

Primary teachers in WA have drawn heavily on *First Steps Language* to support their literacy teaching since it was introduced in the 1990s. The mapping of students' progress on to the continuum has built capacity in making judgements about students' progress over time. The revised *First Steps Reading* materials were published in 2004 and professional learning has been available to teachers in government schools. Demand among schools to access this training has been very strong, and in three terms in 2004, nearly 400 teachers completed the revised *First Steps Reading* professional learning 2-day module.

The *K-7 Literacy Net* also contributes to teachers' capacity to make soundly based judgements about students' progress. The *Net* provides key checkpoints of behaviours that are expected at regular intervals in a student's schooling, with links to English learning area outcomes in the Curriculum Framework. It has been used extensively as a fine-grained and user-friendly tool to help teachers monitor progress and inform evidence based planning, for example, as a target-setting tool in *Getting it Right* schools.

Approaches to Consistency

Curriculum planning

Professional learning modules have been developed to provide teachers with student work samples as profiles of students who have reached the Curriculum Framework Standards, tools and proformas for use as part of the moderation process, quality tasks that can be incorporated into teaching programs and that model effective assessment practice, advice papers on effective practice in relation to policy implementation issues, and authentic exemplars of teachers' planning for learning.

This approach offers teachers a strong suite of resources to build their capacity to design assessment tasks and programs that give students ample opportunities to demonstrate achievement.

Evidence of achievement

School assessment materials have been provided by the Department of Education and Training for use in random sample assessment, but also for schools to use as standardised assessments in single classes and across whole schools. They are designed to assist teachers to assess their students consistently and provide samples for moderation, but also to build the capacity of teachers to design their own quality assessment tasks which are linked to the standards and outcomes framework. The materials are designed for students in Years 3, 5 and 10 so that comparisons can be made between the results of the assessments done in class with those from the random sample used for state-wide reporting purposes. These resources are provided free to Western Australian schools.

Validation of teachers' judgements

Moderation has been a key feature of the *Curriculum, Assessment and Reporting K-10 Policy and Guidelines* document. Procedure 3.3 requires all public schools to implement system-endorsed moderation processes to promote consistent teacher judgements. Teachers engage in ongoing moderation activities within and between schools using moderation processes and support materials.

System-endorsed moderation has two components: a school component for internal consistency managed at school level; and a system component for system comparability managed at district level. District offices contribute expert facilitators and coordinated between-school moderation sessions. Staff working in the district office are regarded as the administration or policy experts and teachers are seen as the content experts who contribute that knowledge to formal moderation processes.

In 2004-2005 all teachers in Years 3, 5 and 7 attended Making Consistent Judgement professional days. This was a major state-wide initiative and teacher participation in the program was fully funded with a significant budget. Over the course of three days teachers were given professional development on Phase 2 of the Curriculum Improvement Program, developing a shared understanding of the English or Mathematics achievement targets, outcome-based assessment and the need for sufficient, 'quality' evidence of student achievement. Collaborative planning for the

collection of work samples demonstrating achievement was completed at the end of the first day in order that teachers could bring appropriate materials on the second day to explore system-endorsed moderation using those work samples. Using the moderation process and system assessment data to make consistent judgements formed the bulk of the work undertaken by teachers on the third day of the professional development scheme.

Western Australia has developed a moderation protocol incorporating the following norms:

- adopt a sense of responsibility for the group
- attend to others and listen
- cooperate in good faith
- aim for consensus decision making
- confront problems respectfully
- allow and give no put downs
- accept where others are at
- suspend judgement.

In moderation there are specific roles for participants. These include the teacher presenter who presents the piece of student work being moderated, the teacher moderator, the facilitator and the content expert. The moderation protocol and arrangements are derived from Kathleen Cushman (1996) “Looking Collaboratively at Student Work: An Essential Toolkit” and were incorporated into the *Making Consistent Judgements* professional reading.

In 2006 the Online Professional Learning pilot project provided teachers from geographically dispersed schools in the Pilbara region of WA with the opportunity to participate in *Making Consistent Judgements* professional development in an online environment. It is planned that this approach will be implemented across the state.

Evidence of Improvement

The Western Australian Department has not systematically analysed teacher judgement results. However, as with other states, there is anecdotal evidence of improved consistency of teacher judgement as a result of these initiatives.

There is scope for comparing teacher judgements with the results of the random sample external assessments that occur annually however this is not currently done.

Summary

Consistency of teacher judgement has been a priority in Western Australia, and the 2005-2006 state-wide program of *Making Consistent Judgements* professional development workshops provided a significant opportunity to develop teachers’ skills and knowledge. Other initiatives, such as the *Getting it Right Literacy and Numeracy Strategy* and the *First Steps* in Mathematics and Reading, aligned with the Curriculum Framework, have strengthened assessment practices in schools.

5. INTERNATIONAL PERSPECTIVES

Three school systems were selected for a review of international approaches: New Zealand, England, and the province of Ontario in Canada. In addition, in reviewing the research and evidence base for approaches to improving the consistency of teacher judgements against standards, some reviews of practices in a number of different countries were located (see Chapter 3).

A summary of the standards frameworks and key features of the approaches to developing consistency of teacher judgement in the three international jurisdictions is provided in Table 3.

The same organising framework as was used for the descriptions of Australian approaches has been used in reporting the investigation of international approaches:

- Teachers' knowledge and understanding of the standards
- Curriculum planning; opportunities for students to demonstrate achievement of standards
- Evidence of student achievement
- Assessment of the evidence against the standards
- Validation of teachers' judgements.

Table 3: Standards frameworks and key resources for developing consistency of teacher judgement in three international jurisdictions

	Standards Framework	Key Resources
New Zealand	<ul style="list-style-type: none"> • New Zealand Curriculum Framework 	<ul style="list-style-type: none"> • Assessment Resource Bank • Assessment Tools for Teaching and Learning (asTTle) • Exemplars (online & provided to schools) http://www.tki.org.nz/r/assessment/exemplars/ • Professional learning • Online moderation support
Ontario, Canada	<ul style="list-style-type: none"> • The Ontario Curriculum (Elementary & Secondary) 	<ul style="list-style-type: none"> • Program Planning and Assessment documents • Council of Ontario Directors of Education advice • CODE templates and tools • CODE moderation advice • Ministry of Education annotated work samples • Assessment Companion
England	<ul style="list-style-type: none"> • National Curriculum 	<ul style="list-style-type: none"> • Local Authority visitation moderation • Training and advice • National Curriculum in Action website http://www.ncaction.org.uk/ • Building A Picture of What Children Can Do pack • Assessment and Reporting Arrangements Handbooks

NEW ZEALAND

Context

Education in New Zealand is a national system, and schools operate under the New Zealand Curriculum Framework. The Framework is being revised and a draft of the new curriculum is currently being reviewed by stakeholders who have been invited to comment.

The Curriculum Framework in its current form covers students from Year 1 to Year 13 of state schooling. It specifies eight essential learning areas and incorporates eight essential skills to be developed across the curriculum. The draft new curriculum outlines eight learning areas and five key competencies. In each learning area there are national curriculum statements which specify the learning outcomes for all students. In each statement, several strands of learning are identified, each with one or more achievement aims. For each of these strands, there are sets of specific objectives, referred to as the achievement objectives. These objectives are set out in a number of levels, usually eight, to indicate progression and continuity of learning throughout schooling from Year 1 to Year 13.

Part of the National Assessment Strategy begun in 1999, and to be reviewed in 2007 to better align with the new curriculum framework, was a focus on the development of key assessment resources for teachers and schools. These resources include the Assessment Resource Bank (ARB) and Assessment Tools for Teaching and Learning (*asTTle*). Part of the documentation for *asTTle* includes an online professional development workshop on social moderation.

A number of strategies and policy and projects related initiatives are accompanied by professional development which provides teachers with the tools to achieve consistent judgements. The NZ Ministry of Education also encourages teachers and schools to participate in Assess to Learn (AtoL) projects and to participate in internal and cross-school assessment practices. The AtoL programme provides primary and secondary schools with facilitators over a two to three year period to develop assessment policies and procedures. The facilitators work with teachers to increase their knowledge of assessment and link it with their teaching practice.

While there is no formal approach to improving consistency of teacher judgement in NZ a number of quality assessment resources have been developed including assessment tasks and tools, curriculum exemplars and teacher professional learning materials. The focus is on providing support at a school policy and implementation level for establishing assessment practices in schools.

Approaches to Consistency

Supporting knowledge of the standards

A review of the current curriculum framework and introduction of the revised version means that teachers will have new material with which to become familiar.

It has been observed that a major issue facing NZ schools is that teachers do not have a common conception of student learning progress, and there is a need to produce curricula documents to document progress and show what it looks like in different areas of learning.

Evidence of achievement

asTTle is a tool for assessing literacy and numeracy in both English and Māori developed by the University of Auckland and provides teachers, students, and parents with information about a student's level of achievement, relative to the curriculum achievement outcomes, for levels 2 to 6

and national norms of performance for students in Years 4 to 12. It is an electronic resource on CD-ROM (an online version is currently in development and trial). The software generates tests from a pool of questions according to the content and difficulty levels the teacher requires. Students' answers to questions are input and scores are generated and can be compared through comprehensive reporting with nation-wide norms.

The report generated once a student has sat through an *asTTle* assessment shows what a student knows, what gaps are displayed in their learning and what they need to learn next. Comparison with other students nation-wide is also a key feature of this tool.

The Assessment Resource Banks (ARBs) consist of over 3000 curriculum-based assessment resources designed for students working at levels 2-5 for use in NZ schools. Resources are available for English, Mathematics and Science strands. The Bank is only accessible to staff in NZ educational institutions or students in teacher training in NZ, however some sample resources are available online. Each resource is linked to a strand, objective and level from the Curriculum Framework and is accompanied by a marking guide.

Assessment of evidence against the standards

Exemplars of student work and assessment are widely available, and professional learning in and around assessment is well established. The exemplars assist teachers to make consistent judgements of student achievement against the NZ Curriculum Framework standards.

Exemplars in the learning areas for English, mathematics, the arts, science, technology, health & physical education, social studies, te reo Māori, pāngarau (mathematics), pūtaiao (science), and hangarau (technology) have been provided to schools and published online at the Te Kete Ipurangi website. The exemplars are all authentic pieces of student work annotated to illustrate learning, achievement, and quality in relation to the levels in the national curriculum statements.

Validation of teachers' judgements

Social moderation to validate teacher judgements is supported online but there is no indication of formal, systemic moderation systems in NZ.

Evidence of Improvement

As in most other jurisdictions there is no formal evidence of improved consistency in teacher judgements in NZ. Speaking of the strategies employed in NZ to improve consistency, one of the experts consulted in the investigation commented that

While all of these strategies help teachers to form views about the expected levels of achievement, in and of themselves of course they don't guarantee consistency of teacher judgement. That comes about not just from having available useful curriculum and assessment resources but through professional dialogue, as part of either in-school or national professional development (which generally takes place shortly after the implementation of a new policy or programme).

Summary

The *asTTle* (Assessment Tools for Teaching and Learning) is recognised as a high quality assessment tool valued by teachers because of the rich interpretation it provides of student performance. It has a strong empirical and theoretical research base. It is based on a rationale for using assessment to improve the quality of teachers' instruction and students' learning. The *asTTle* allows teachers to choose the content and difficulty level of a test to use in the classroom,

and generates a range of interpretative reports. It is available on CD-ROM, free to all schools. (www.asttle.org.nz)

* * * * *

ONTARIO

Context

The Ontario curriculum framework is divided between Elementary, Years 1 to 8, and Secondary, Years 9 to 12. For each subject a curriculum document outlines the curriculum expectations and achievement levels for each year of schooling in that subject. In addition, advice about planning for learning and assessment is contained within the curriculum document for each subject alongside the roles and responsibilities of students, teachers, parents and caregivers, principals and the community in the learning associated with each subject area. There are curriculum documents for the arts, French as a second language, health and physical education, kindergarten, language, mathematics, native languages, science and technology and social studies in the elementary area, while for secondary students a number of additional subject areas are included. Explicit curriculum documents for program planning and assessment for Years 9 to 12 which include requirements for consistent judgement also form part of the Ontario curriculum documentation.

Approaches to Consistency

Curriculum planning

Collaborative planning and curriculum analysis exercises are suggested by the Council of Ontario Directors of Education (CODE) to enhance teachers' planning processes and curriculum knowledge. Indeed collaborative planning is described in the document as the single greatest way to assure consistent practices and judgements.

Evidence of achievement

Grades 9 to 12: Program Planning and Assessment, part of the Ontario secondary school curriculum, makes clear the expectation that teachers when assessing student work are assessing against the provincial curriculum expectations and achievement levels. It also defines assessment as the process of gathering evidence from a range of sources which accurately reflect student achievement of curriculum expectations. However, there is limited assistance for teachers in terms of generating and collecting appropriate work samples to build the body of evidence of achievement of the standards.

The Council of Ontario Directors of Education (CODE) provide some support with advice on developing quality assessment tasks which incorporates assessment design templates and tools, for example rubrics, checklists and exemplars (Council of Ontario Directors of Education, May 2006). The emphasis is on providing students with the opportunity to demonstrate achievement of the standards through quality tasks.

Assessment of evidence against the standards

To help teachers make judgements of student work against the curriculum expectations, exemplars of student work aligned to the curriculum are available through the Ministry of Education website. This assists teachers to assess to the standards and identify the range of student achievement through comparison with examples of student work at different levels.

Validation of teachers' judgements

Moderation processes are also described by CODE as suggested activities to enhance consistency and includes shared marking of student work, collaborative planning of common assessment activities and collaborative use of rubrics, exemplars and sample answers.

School leaders have access to CODE professional learning opportunities which focus on the role of school leaders in the assessment and monitoring process. The Ministry of Education has also produced an Assessment Companion to assist teachers in curriculum unit planning which outlines the many ways in which students can demonstrate their achievement and methods by which this achievement can be recorded. While not explicitly geared towards consistency it addresses the concepts of curriculum planning and evidence of achievement which are essential to consistent judgement.

Students in Ontario participate in province-wide assessments at primary, junior and secondary levels for the purposes of improving education as the individual, school and provincial levels. However, data from these assessments is not formally compared to the in-school teacher judgements as a measure of consistency.

Evidence of Improvement

Curriculum analysis exercises and collaborative planning in Ontario help teachers to build their knowledge and understanding of the standards against which they assess their students in the province. Program planning which provides students with an opportunity to demonstrate achievement of standards is highlighted by the curriculum documents themselves and supported by additional materials produced by CODE and the Ministry of Education. The Ministry of Education supports teachers through online exemplars and work samples. This assists teachers to make their own assessments against the standards. Notions of moderation to validate teacher judgements are not as well developed as the assessment planning and curriculum knowledge key elements. However, social moderation, collaborative planning of common assessment tasks and collaborative assessment are all mentioned as methods for ensuring consistent teacher judgement against the Ontario curriculum standards.

* * * * *

ENGLAND

Context

The National Curriculum guides the work of teachers in England. The curriculum consists of Four Key Stages aligned to age and year levels. Key Stage 1 is for children 5-7 years of age in Years 1-2 of primary school. Key stage 2 is for children in years 3-6, key stage 3 covers students from Years 7-9 and key stage 4 is for students from Years 10-11. At each Key stage there are a number of compulsory subjects and each subject at each key stage is accompanied by a programme of study which provides the basis of teacher planning. In addition, attainment targets set out the knowledge, skills and understanding which pupils of different abilities and maturities are expected to have by the end of each key stage.

Attainment targets consist of eight level descriptions according to difficulty and one description for exceptional performance above level 8. Each level describes the types of and range of performance that students working at that level should demonstrate. When assigning a level of attainment to a student at the end of a key stage, teachers judge which description best fits the student's performance. Level descriptions are not intended to be used to assess individual pieces of work, rather teachers use them to assess students' performance over time in a range of contexts.

Statutory assessment at the end of each key stage has been implemented. Teachers are required to report annually on students' progress across the key stage. Schools submit teacher assessment levels to their Local Authorities according to its instructions and the assessments are in turn submitted to the Department for Education and Skills (DfES). At the end of each of the Key Stages a combination of teacher assessment and standardised testing results is used to assess the attainment levels of students.

Approaches to Consistency

Assessment of evidence against the standards

The National Curriculum in Action website hosts annotated work samples at each Key Stage and at each level. The samples are accompanied by an extensive commentary.

The Building a Picture of What Children Can Do pack contains a booklet, CD-ROM and video designed to assist in the Key Stage 1 teacher assessments; it was distributed to all schools in 2005.

Validation of teachers' judgements

Schools are required to be involved in a formal Local Authority moderation process at least once every four years for the end of Key Stage 1 as one part of the continuous moderation process. This moderation process follows a visitation model where experts visit schools and classrooms to moderate their practice. Local Authorities are required to identify 25 per cent of their schools for such formal moderation each year. Each Local Authority sets up moderation arrangements that are appropriate for their area. The National Assessment Agency (NAA) provides Local Authorities with guidelines and examples of good practice. It is recommended by the NAA that the 75 per cent of schools that are not being formally moderated in the current year take part in inter-school social moderation activities. Moderation arrangements are governed by the Assessment and Reporting Arrangements which are reviewed and republished annually by the Qualifications and Curriculum Authority (QCA).

Local Authorities develop their own moderation processes with the goal that schools and teachers within their jurisdiction develop clear understandings of the national standards and apply those standards consistently (National Assessment Agency, 2004). As noted above, this is a visitation moderation model. For example, in Dorset the moderation team is lead by a recently retired head teacher, supported by four practising Year 2 teachers and one experienced Year 2 teacher working in Year 1, and a team appointed by the Local Authority including two head teachers. The moderation team is trained in the process of teacher assessment, arrangements for the moderation process and agreement trialling which sets the standards and benchmarks for students' work.

The moderation process is made up of three parts. The first, agreement trialling, takes place across the jurisdiction and is attended by Year 2 and Year 3 teachers. The material used for agreement trialling is drawn from writing samples covering the range from working towards level 1 to level 3, mathematics tasks from children, the level or grading of which is set prior to the meeting. These items of student work are used as exemplars for determining levels of work during the meeting. During agreement trialling the moderating team works alongside teachers as supports and facilitators.

The second part of the moderation process involves consensus moderation meetings conducted across the jurisdiction in twilight or half day sessions. Staff are released from school to attend these meetings and replacement teachers employed to cover their absence. Each meeting is attended by the lead moderator, two members of the moderation team and the Senior Inspector for Assessment. Teachers are asked to bring work samples from the level boundaries for writing,

reading and mathematics and the samples are moderated in small groups. Additional support is available through this process for newly qualified teachers and teachers who are new to Year 2.

The final stage in the process is the moderation meetings which occur in 25 per cent of identified schools each year and any other schools which are of concern to the Local Authority. Visits by the moderation team occur in afternoons by prior agreement with the school head teacher and are focussed on writing. For the visit a teacher is asked to select three children, one working at average standard for the class, one working below and one working above. The writing of these children is examined and discussed with the children concerned. Oral feedback is provided to the teacher, the head teacher and the assessment coordinator at the end of the visit. This is just one example of Local Authority moderation processes (in Dorset) and they vary from area to area.

In addition to these formal moderation processes Local Authorities are encouraged to develop additional support for teachers in their jurisdiction as part of their moderation systems. They are expected to offer schools training and advice on all aspects of Key Stage 1 assessment, ensure that schools have the systems in place to record and process assessment data, and where possible provide moderation support and assistance for schools not included in the mandated 25 per cent group of schools each year.

The Assessment and Reporting Arrangements Handbooks (Qualifications and Curriculum Authority, 2006) provide information for teachers and schools about the assessing and reporting regime, including moderation for Key Stage 1 on an annual basis. The 2007 Handbook for Foundation stage, Years 1 and 2 is available from the QCA website.

At Key Stage 1 reporting on student achievement is done by combining task and test assessments with teacher assessments. Where there is discrepancy between the assessment done by the teacher and the standardised test, the teacher judgement is used. There is recognition that student levels can move before and after the testing period and these students achieve results through regular classroom assessment tasks which are different from those achieved in the standardised tests. At other Key Stages teacher assessments are reported alongside standardised test results; however, at this stage there is no formal moderation process established at Key Stages 3-4.

The *Key Stage 1 Evaluation of Moderation Arrangements* found that teachers reported high levels of confidence when making judgements and kept students' work to support those judgements. However, there was wide variation in the quality of work chosen to be kept and how much of it was kept (Reed & Lewis, 2005). A survey conducted as part of that evaluation asked teachers to comment on how being moderated affected how they made their judgements. Around 150 teachers responded and the most common response was that moderation has confirmed their judgement and indicating that teachers being moderated in that year were making consistent judgements about student achievement at Key Stage 1.

Summary

The English system has a greater emphasis on external evaluation and assessment than its Australian counterparts. Teacher judgements, while important, appear to play a secondary role to statutory assessment at the end of each Key Stage. This is being slowly changed to report both teacher judgements on student learning and statutory assessments, giving more weight to classroom judgements than in the past. Teachers are supported to make assessments against the National Curriculum through access to annotated student work samples.

Use of classroom assessment alongside common tasks and tests helps teachers build up the body of evidence from which they make their judgements about student achievement. Validation of teacher judgement is achieved through a statutory visitation moderation program to one quarter of the schools each year. Teachers who are not involved in the mandatory moderation process are

encouraged to participate in social moderation but this is supported to varying degrees by the different Local Authorities.

6. FINDINGS

A key finding of this investigation is that in all Australian states and territories there are current initiatives, resources and programs designed to improve the consistency of teachers' judgements of students' progress against common standards. This activity reflects a strong interest related to the development and continuing implementation of standards-based curriculum frameworks. We found considerable commonality in initiatives and strategies designed to improve the consistency of teacher judgements.

In order to draw out the commonalities and differences, and the relative effectiveness of these initiatives and approaches, we reviewed the extent to which five key elements of effective assessment systems were incorporated into the approaches investigated. These elements have been identified as underpinning a wide range of assessment policies and practices reported in the research literature (Matters, 2004). The set of key elements is as follows:

- teachers' knowledge and understanding of the standards
- curriculum planning providing opportunities for students to demonstrate achievement of standards
- evidence of student achievement
- assessment of the evidence against the standards
- validation of teachers' judgements.

These elements match the implications for improving teachers' summative assessment identified in the systematic review undertaken by the EPPI-Centre (Harlen, 2004). There is much congruence between the elements described above and the implications for policy and practice identified by that review, including:

- *There is a need for resources to be put into identifying detailed criteria that are linked to learning goals, not specially designed assessment tasks. This will support teachers' understanding of the learning goals and may make it possible to equate the curriculum with the assessment tasks.*
- *It is important to provide professional development for teachers in undertaking assessment for different purposes that address the known shortcomings of teacher assessment.*
- *The process of moderation should be seen as an important means of developing teachers' understanding of learning goals and related assessment criteria.*
- *There should be wider recognition that clarity about learning goals is needed for dependable assessment by teachers.*
- *Schools should take action to ensure that the benefits of improving the dependability of the assessment by teachers is sustained (e.g. by protecting time for planning, assessment, in-school moderation etc).*
- *Schools should develop an 'assessment culture' in which assessment is discussed constructively and positively and not seen as a necessary chore (or evil.) (Harlen 2004, pp 96-97.)*

Overall, in the Australian context, and in international contexts, we found evidence of these elements, to varying degrees, in all of the approaches investigated, including state-wide standards,

a variety of print and web-based resources to support teachers' use of the standards, reporting requirements, and a range of professional learning and moderation activities varying from in-school to state-wide activities.

The Australian and international experience indicates the significance of a comprehensive approach to improving the consistency of teachers' judgements. No single element, no matter how well designed, will be sufficient on its own. There is a need for a range of elements that complement and reinforce each other. None of the systems we examined could be considered as 'best practice' as a whole, as each was incomplete in different ways. However, in each of the systems there were some elements that can be considered 'best practice'.

The approaches which represent the most effective and strongest measures in place that support each of the five elements are outlined in this chapter. Table 4 acts as a guide to these approaches. The table also includes a sixth, more over-arching element of a comprehensive strategy: the extent to which there is systematic monitoring of the take-up of the various approaches and evaluation of their impact.

Table 4: Key elements in a comprehensive strategy to improve consistency of teacher judgements

Purpose	Contexts	Present in Victoria	Examples of effective approaches
1. Building teachers' knowledge and understanding of assessment of, as and for learning	Professional resources	<u>Assessment advice</u>	<p><u>South Australia</u> <i>Let's Talk Assessment ...</i> This newsletter, sent to schools quarterly and available online, examines assessment of, as and for learning from a research and practical perspective and is intended to keep assessment issues and practices at the forefront of teachers' minds.</p>
	Use of research evidence		
	Professional learning activities (in and beyond schools)		<p><u>Multiple States/Jurisdictions</u> Teacher pre-service education Teacher registration On-going teacher professional learning</p> <p>The EPPI-Centre systematic review noted the importance of professional development for teachers in undertaking assessment for different purposes.</p>
2. Developing common understanding of the standards	Annotated work samples	<u>VCAA annotated samples</u>	<p><u>England</u> <i>National Curriculum in Action</i> The National Curriculum in Action website hosts annotated work samples at each Key Stage and at each level. http://www.ncaction.org.uk/</p>
			<p><u>New South Wales</u> <i>Assessment Resource Centre</i> Provides detailed analysis of student work samples assessed against the syllabuses in a format that is easy to follow. http://arc.boardofstudies.nsw.edu.au/</p> <p>The EPPI-Centre systematic review noted the importance of supporting teachers' understanding of the learning goals, and the importance of clarity about learning goals needed for dependable assessment by teachers.</p>

3. Designing assessment tasks	Resources: models and sample tasks	<u>VCAA and DoE</u>	<p><u>Tasmania</u> <i>Guiding Assessment Tasks</i> Guiding assessment tasks are common tasks developed collaboratively by teachers, School Education Division and Office for Educational Review staff. They address the calibrated key elements and are administered to entire cohorts and samples of students and reported back to schools and centrally through the Student Assessment and Reporting Information System (SARIS).</p> <p><u>New Zealand</u> <i>asTTle</i> Software generates tests from a pool of questions according to the content and difficulty levels the teacher requires. Students' answers to questions are input and scores are generated and can be compared through comprehensive reporting with nation-wide norms.</p> <p><u>New Zealand</u> <i>Assessment Resource Bank</i> Collection of Standards aligned assessment tasks for use in classrooms consisting of tasks for students to complete including selected and constructed responses, practical performance and pen and paper responses and oral presentations. They are accompanied by teacher information pages and marking guides.</p>
4. Assessing evidence against the standards	VCAA Assessment Maps		<p>New South Wales <u>Assessment Resource Centre Website: authentic exemplars of student work aligned to A-E grade levels.</u></p> <p>South Australia <u>Assessing SACSA Outcomes Through Standards. CD-ROM with exemplars, plan for teaching and learning, student evidence, commentary.</u></p>
5. Validating teacher judgement	Collaborative assessment meetings in schools & networks	<u>Individual schools and networks.</u>	<p><u>Tasmania</u> <i>QMAP</i> Three student free professional learning days, two facilitated by the department and one for schools to use for internal school level moderation. Department days run through school clusters and focus on collaboratively assessing student work using an assessment protocol.</p> <p><u>New South Wales</u> <i>Consistent Teacher Judgement Online</i> Online programs and materials targeted at teachers in rural and remote locations.</p> <p><u>Queensland</u> <i>New Basics Moderation</i> Incorporates the professional learning aspects teachers value when they are able to discuss student work with their peers with a level of oversight provided by the moderation teams and submission of grades for ratification and eventual reporting to the central body.</p> <p><u>South Australia & Tasmania</u></p>
	Resources		

SACSA Moderation Protocol/ Tasmanian Essential Learnings Assessment Protocol

Developed with reference to the 'Tuning Protocol' to provide structure for the management of conversation around student work that occur during moderation.

The EPPI-Centre review emphasised that the process of moderation should be seen as an important means of developing teachers' understanding of learning goals and related assessment criteria, and that schools should take action to ensure that the benefits of improving the dependability of the assessment by teachers is sustained (e.g. by protecting time for planning, assessment, in-school moderation etc).

6. Monitoring take-up and evaluating impact

Collecting and analysing data on changes in teacher practice, the extent to which teacher judgements are becoming more consistent, and the costs of the initiatives

Some school and cluster-initiated data and evaluation

Comparison of teacher judgement data with AIM data in English and Mathematics

Tasmania

Quality Moderation of Assessment Process (QMAP)

A system-wide quality control mechanism and feedback for teachers is provided through the Student Assessment and Reporting System (SARIS); enables modelling of teacher judgements school data from state-wide testing Evaluation of the state-wide professional learning days and moderation exercises through feedback from participants

Western Australia

England

Schools are required to be involved in a formal moderation process every 4 years. This involves a school visit by experienced school leaders and teachers, review of school policies and practices, and analysis of student work and teacher assessments

Key Stage 1 Evaluation of Moderation Arrangements

Included teacher surveys about how they make judgements, record keeping,, and how involvement in moderation, affected their approach

New Zealand

The National Assessment Strategy, which began in 1999, is undergoing a major review in 2007

A longitudinal approach is the only appropriate way of gathering evidence of the consistency of judgements over time

The components of some effective approaches selected for further analysis here were selected on the basis of meeting one or more of the elements of best practice in developing consistent teacher judgement. Additionally, practicality of implementation, overall effectiveness of design, state-wide wide implementation, sustainability and focus on consistency, were important considerations in the selection process.

Curriculum Planning and Effective Assessment Practice

Limited evidence was found of resources and practices for curriculum planning directly linking standards and assessment tasks in order to address issues relating to consistency of teacher judgement. For example, the South Australian newsletter *Let's Talk Assessment* is one of the few resources which looks at overall effective assessment practices, is distributed state-wide and targets teachers directly. The newsletter is distributed to schools on a quarterly basis; however, copies are also available online from the SACSA website so it is accessible to all teachers in the South Australian education system. The style of the newsletter mixes research evidence with key quotes and suggestions from the Assessment Reform Group, and advice for assessment practices and strategies as well as practical examples of how assessment can be done in the classroom and this is one of its particular strengths.

This approach to tackling curriculum advice and highlighting effective assessment practice is relatively inexpensive. The benefits are extensive in that the newsletter reaches all schools and is likely to be read widely, effectively reaching its audience. It provides stimulus for extended discussion of assessment practice in schools and keeps that discussion current and regular.

Teachers' knowledge and understanding of and assessment against the standards

High quality, benchmarked, annotated work samples provide an effective method for building a common understanding of what work looks like at given levels and points in a Standard.

The *National Curriculum in Action* (England) and *Assessment Resource Centre* (NSW) represent useful examples of comprehensive collections of work samples. They are complete catalogues which are easily navigable and practical resources to assist teachers assess their students' work. The student work samples are authentic, provided by teachers in the system and indicative of what teachers can expect to see produced in their own classrooms. Quality is maintained by careful selection of the samples. In the case of the NSW Assessment Resource Centre samples, grading judgements about the samples are made by practising teachers considering the grade descriptions, the syllabus outcomes, and content for the stage and the Foundation Statements. Both collections contain work that is representative of what is produced by students at a range of levels of achievement.

A potential difficulty with online materials such as these is that there is no system-level check on whether teachers are using the materials to enhance their own assessments. Such resources need ongoing promotion and the ways that they improve teachers' work need to be continually reinforced.

The range and number of these resources means that online delivery is the most cost-effective way of providing access for all teachers, and keeping materials up-to-date. A further benefit of this kind of resource is that it provides a large collection of assessment tasks which teachers can adapt for their own purposes as well as a more general picture of what student work at any particular level or stage looks like. The samples are current and that the bank is refreshed with new samples on an ongoing basis.

Evidence of student achievement

There was considerable diversity in the methods for obtaining evidence of student achievement across the jurisdictions. One effective way of ensuring consistency is to build a bank of common assessment tasks which teachers use to assess the achievement of their students, backed by moderation and analysis of the results. The Tasmanian Guiding Assessment Tasks and the New Zealand *asTTle* program represent two models of common assessment task management and administration methods which provide a consistent assessment regime and incorporate feedback. Importantly, these assessments do not incorporate teacher judgements, but provide an external reference point for teachers when they come to make their own judgements about student achievement in other assessment tasks.

The Guiding Assessment Tasks in Tasmania are administered to students and marked centrally then incorporated into the SARIS database. These externally assessed work samples contribute to the body of evidence from which teachers make judgements and provide a benchmark for teachers to use when assessing their students' work as a whole. State-wide whole cohort and sample testing and marking of tasks such as these requires extensive resources in terms of administration both centrally and in schools, and staffing for marking and analysis of results. The benefits are in the way in which the tasks act as a robust quality control mechanism providing a centralised check on what teachers report about student achievement through the SARIS system. The results form an example of student achievement against the Standards in particular domains at a given point in time and are intended to provide the basis for internal moderation processes and a reference for other teacher assessed work.

In NZ the *asTTle* program provides a cutting-edge, technology driven and accessible means for teachers to build assessments for learning that are linked to Standards and provide comparable data on student achievement nation-wide. Teachers are able to select the level of assessment and use the assessment tool at whatever stage they wish to gain feedback on student achievement and where gaps in learning might exist. The tool helps teachers select items for testing, then scores and analyses them before providing links to appropriate teaching resources. The software creates tests from an item and test data bank incorporating closed and open-ended items mapped to curriculum statements. Teachers print the test and scoring key, administer the test and then enter the score data. Six major reports incorporating analysis, further action and available resources can be produced.

The cost of establishing such a system is necessarily high. Development and calibration of the item bank and development of the software and online environment requires specialist skills and regular monitoring and revision. However, the use of the software in schools is efficient and the flexibility certainly of benefit to teachers who have access to the tests at times which are convenient to them and require less planning time than the preparation of traditional assessment tasks. The online delivery means that ongoing costs associated with using the software are small. The management of test items to avoid repetition means that the resource is sustainable and can be used multiple times.

The samples in the NZ Assessment Resource Bank were developed by the NZ Council for Educational Research after trialling and advice from curriculum experts and teachers and are regularly revised and reviewed to maintain currency and sustainability of the resource. There is a substantial development cost associated with this process. High quality examples of assessment tasks provide a strong model from which teachers can build their own assessment tools, and tasks and marking guides show how student work might represent achievement of the Standards identified for the task. In NZ the bank is available online and, as with the annotated work samples discussed above, this represents a practical solution to providing an extensive and comprehensive resource which is accessible to teachers across the system. Another strength of this kind of

resource is the flexibility it offers teachers and schools in what they want to assess, contrasting with the more rigid nature of the common assessments tasks like the Guiding Assessment Tasks.

Validation of teachers' judgements: moderation

It is clear that the key to validating teachers' judgements is a systematic, organised program of moderation activities undertaken by teachers within their own schools and between schools. Moderation not only ensures consistent judgement when assessing student work but helps build teachers' knowledge and common understandings of the standards and assessment tasks which enable students to demonstrate achievement of the Standards.

Three different approaches examined during this investigation illustrate effective practice in moderation. In this context, moderation involves processes where teachers work with colleagues to collaboratively examine students' work, and reach consensus on an assessment of the work against common standards.

The Tasmanian approach, located within the overall Quality Moderation of Assessment Process (QMAP) from 2005-2006, brought teachers together in clusters at the same time for state-wide moderation of student work. Using a specified protocol for the discussion of work, teachers were trained in moderation practices and processes and developed skills in collaborative assessment. This moderation program required significant central and regional staffing resources and teacher time, including three student-free days per year, and therefore represents a costly approach to moderation. However, evidence gathered by the Department via survey and feedback forms indicated that teachers felt more confident about the requirements of the Tasmanian Standards and that their ability to assess their students' work consistently had improved, thus achieving the objectives of the moderation process.

Use of a structured protocol, grounded in the work done by the Coalition for Essential Schools for conducting the discussion was a key feature of this moderation program. This protocol, adapted by a number of jurisdictions investigated for this report, was an element identified by participants in the moderation process which ensured that the process of moderation remained focussed and that teacher presenters and participants maintained respect for the professional judgements being examined and reviewed.

The state-wide nature of this moderation ensured that teachers across Tasmania were participating in the same process. While logistically challenging, this kind of state-wide moderation and training was possible given the relatively small size of Tasmania Department and the staff resources dedicated to the program.

The NSW Consistent Teacher Judgement Online program addressed the logistical difficulty associated with moderation in rural and regional areas. The online environment allowed teachers to examine student work with colleagues despite geographical separation. Online delivery removes expense and time associated with travelling to participate in moderation activities. Teachers in different locations in a large state who might otherwise have no opportunity to discuss assessment and plan for consistency with colleagues were able to meet regularly via internet technology and conduct moderation in much the same way as it would occur in a face-to-face meeting.

The technological requirements of the program -- the web board meeting environment and the capacity to make digital copies of student work and upload them to the shared space -- posed potential challenges. However, appropriate support mechanisms, including links to software downloads and advice, supported users of the NSW system. This kind of support is essential for any large scale implementation of this kind of model. The use of trained e-moderation facilitators to direct the online moderation process added a further layer of assistance to users. The main

benefit in facilitating this kind of moderation is that teachers have the opportunity to improve their understanding and application of the standards through collegial conversation.

The Queensland New Basics approach to moderation whereby teacher judgements were moderated internally and then ratified by an expert team combines the benefits of inter- and intra-school moderation with a central, over-arching quality control mechanism. While the ratification process was important as a checking mechanism for teacher judgements, the expert team acted primarily as moderation facilitators, meaning that the weight of responsibility for awarding grades and maintaining consistency remained with teachers. In addition to checking the moderated student work at the end of the assessment process, expert moderators were available to provide advice to teachers and schools about the implementation of their internal moderation processes. This brought the benefit of improving and standardising the moderation activities within schools and achieving consistency between schools because the advice and support given was uniform across the schools participating in the New Basics Trial.

This process of ratification is unique in Australia, although it is similar in some ways to the moderation model used in England at Key Stage 1. It is the only model where moderation activities and teacher judgements are checked for effectiveness, reliability and accuracy by expert teams.

Monitoring take-up and evaluating impact

The Australian and international experience reviewed in Chapters 4 and 5 provided few instances of evaluations of the initiatives intended to improve the consistency of teachers' judgements. As the summary provided in the final section of Table 4 indicates, some interesting approaches to evaluation can be found in England, New Zealand and Tasmania. However, even in these cases the evaluations that have been done or are underway tend to focus on just one aspect of the initiatives, and collect only a limited range of evidence, generally in the form of teachers' views on the initiatives. In particular, there seemed to be little documentation of the resources involved in the various approaches, or assessments of whether the benefits justified the costs.

At a minimum, a systematic evaluation strategy would include collecting and analysing data on:

- The extent to which the initiatives are known about and taken up in schools
- The extent to which the initiatives change teacher practice
- The extent to which teacher judgements are becoming more 'accurate' and more consistent
- The conditions that lead to beneficial outcomes, and the factors that inhibit successful implementation
- The resources involved in the initiatives and their costs
- The relative cost-effectiveness of alternative approaches.

In most of these respects, a longitudinal approach will be the only appropriate way of gathering evidence of the consistency of judgements over time.

There would be value in Victoria capitalising on the diversity and variation within its school system by testing new initiatives on a pilot basis, with volunteer schools and regions, before moving to more widespread implementation. Identifying the factors involved in successful

innovations, and creating the conditions for their dissemination, mainstreaming and sustainability in other schools, are central to an effective implementation strategy.

The overall monitoring and evaluation strategy needs to be developed and integrated at the school system level, although it is also important for regional groupings and individual schools to be closely involved in monitoring their own activities. The regional and school-level evaluations have greater value, though, when there is a framework that makes it possible to compare those results with regions and schools in other locations. Victoria already has a strong culture of such approaches through the opportunities schools have for ‘like school’ and state-wide comparisons of school performance on student assessments in literacy and numeracy. It should be possible to use a similar approach to strengthen the knowledge base about what works best in improving the consistency of teachers’ judgements.

Assessments of Cost-Effectiveness

Much educational evaluation focuses only on the outcomes of different programs – whether a particular program leads to higher performance on the outcome measure concerned. While a focus on measuring outcomes is necessary, it is not sufficient for policy development. There is also a need to analyse the resources involved in the program(s) and their costs. The cost information is needed to identify approaches that are *cost-effective*, that is, which achieve particular outcomes with the least use of resources or, alternatively, which enable higher outcomes to be achieved with the same level of resources. The concept of cost-effectiveness brings together evidence on benefits and costs within the same decision-making framework.

As the school system reviews in Chapters 4 and 5 indicated, there is only limited ‘hard’ evidence on the benefits of different approaches to improving the consistency of teachers’ judgements, and a real paucity of data on what these various approaches cost. Thus, assessments about the likely cost-effectiveness of programs need to be approached in cautious terms, as they rely on judgements from patchy data rather than on rigorous evaluations.

Each of the approaches identified in Table 4 had recognised benefits in strengthening teachers’ capacity to make consistent judgements of students’ work against standards. Several examples attest to this. Evidence of increased consistency derived from the professional practices promoted by the various initiatives was found in the observations of the managers of the approaches, for example, the SACSA team in SA, and the QMAP team in Tasmania. Evidence of teachers’ increased confidence came from anecdotal reports, and, in Tasmania, from surveys of teachers following moderation workshops. Evidence from CSF data provided to VCAA by Victorian schools over recent years, and some evidence collected in the SARIS system in Tasmania, also provided some evidence of improved consistency in teachers’ judgements. However, while there was widespread agreement about the value of the approaches in improving consistency, no systematic longitudinal processes for gathering evidence of improved consistency were located.

The benefits to teachers’ professional learning were frequently cited in interviews with assessment program managers throughout Australia. Of particular interest is a key finding from the systematic review of the evidence of reliability and validity of assessment by teachers used for summative purposes conducted by the Assessment and Learning Research Synthesis Group at the EPPI-Centre at the Institute of Education, University of London (Harlen, 2004). The review found that

Moderation through professional collaboration is of benefit to teaching and learning as well as to assessment. Reliable assessment needs designated time for teachers to meet and take advantage of the support that others including assessment advisers can give. (Harlen, 2004, p 84).

Notably, many of the approaches bring considerable benefit in the form of teachers' professional learning, and consequent improvements in teacher assessment. This is a significant finding of this investigation.

The experience in New South Wales, New Zealand and elsewhere is that online delivery of sample assessment materials is a cost-effective means of providing teachers with access to a large bank of exemplar assessment materials and keeping them up-to-date. Online delivery also provides opportunities for teachers and schools to adapt materials to meet their own needs. Although the start-up costs of online systems can be high, the amortisation of costs over such a large school system as Victoria, and over time, mean that the annual costs per teacher are relatively low.

Experience indicates that using teachers' professional learning to build understanding of the curriculum in combination with planning assessment for learning is a beneficial approach. Although the cost of providing teachers with extensive professional learning time is necessarily high in the first instance, it is a critically important means of developing teacher expertise.

A comprehensive strategy also encompasses means of regularly updating teachers' knowledge and improving practice. South Australia's *Let's Talk Assessment* newsletter, which is distributed to schools quarterly and available online, seems to be a relatively low-cost means of keeping assessment issues prominent in schools.

The overall message of the review is that school systems need to take a comprehensive approach to improving the consistency of teachers' judgements. This includes building up teachers' skills and knowledge about curriculum standards and assessment, providing them with exemplar materials, ensuring that there are opportunities for teachers to meet regularly for moderation purposes, monitoring the relationship between teachers' assessments and other data on student performance, and evaluating outcomes and costs on an ongoing basis. No single element will be sufficient on its own, and the various elements are not substitutes for each other. Some may be put in place before others, either because some of the foundations are already there, or the need is more pressing. However, eventually all the elements will need to be there so that they complement and reinforce each other.

Table 5 draws together the review to identify the likely costs and benefits of implementing a comprehensive strategy in Victoria, recognising those elements that are already in place to some extent.

Table 5: Types of resources and illustrative costs in implementation of a comprehensive strategy in Victoria

Elements of approaches to consistency	Aspects not currently present in Victoria	Types of resources and costs	Benefits
Assessment advice materials	Currently available, needs to be kept up-to-date Assessment newsletter	Staff time Production Distribution	Maintains focus on assessment & consistency Targets schools & teachers directly Strengthening of assessment practices in Victorian schools.
Preparation of sample assessment tasks and annotated work samples	Currently available; need for more comprehensive exemplars showing how assessments are	Staff time Consultancy fees Ongoing revision/maintenance Collection of samples	Teacher professional learning Leadership

	linked to the learning goals of the VELs standards		
Common assessment task preparation/generation	No current arrangements similar to the NZ <i>asTTLe</i> resource	IT support Expert consultancy fees	System-wide commonality
Online materials and processes	Material online need to be continually updated, and processes for online moderation need to be developed	IT support Access time Development costs	Easy of accessibility Potential for continual updating to maintain currency
Assessment advisers, or support officers	No statewide team currently in place	Salaries. Training Travel funds Access to resources	Expertise and advice to schools
Teacher selection of work samples	Little systematic focus on this aspect	Staff time Support to schools from assessment advisers	Teacher ownership of assessment process
Teachers' attendance at moderation meetings	Systematic, system-wide approach required	Teacher replacement costs Travel expenses	Teacher professional learning
Collaborative planning time	Systematic, system-wide approach required	Staff time in school	Teacher professional learning
Monitoring and evaluation system	Systematic, system-wide approach required	Central office and regional staff time in developing an evaluation framework successful approaches Collection and analysis of ongoing data, preferably on a longitudinal basis Case studies of particularly effective approaches	Quality assurance Modification of approaches in light of rigorous data on benefits and costs Use of pilot studies to assess cost-effectiveness before mainstreaming Opportunities for schools to monitor their performance against like-school and statewide data

7. CONCLUSIONS AND RECOMMENDATIONS

The main purpose of this investigation was to identify best-practice approaches for improving the consistency of teacher judgements, and to make recommendations for the implementation of a coherent approach linking to and building upon the Victorian Essential Learning Standards (VELS).

A general finding of the investigation is that all Australian states and territories currently operate a range of strategies related to the consistency of teachers' judgements of students' progress against common standards. Some operate at the level of large scale system-wide initiatives, others are more localised, at the school or network level. The evidence suggests that initiatives are needed at both levels. The development of these strategies is related to the development and continuing implementation of standards-based curriculum frameworks. In Victoria the practice of reporting students' progress against the *Curriculum and Standards Framework*, and more recently the *Victorian Essential Learning Standards*, has been in place for several years, as has been the practice of the collection by the Department of aggregated data in relation to students' achievement in English and Mathematics.

The standards constitute clear maps of the levels of achievement in areas of learning, and provide a means of planning teaching and learning activities, and also of monitoring students' progress across the years of schooling. The practice of reporting students' progress against the standards is continually developing. While the results of state-wide testing programs can be, and are, reported against the standards, teachers have continuous access to a rich array of evidence of students' work and achievement in classrooms, and there is great value in judging this evidence against the standards and reporting these judgements. Consequently, there is increasing recognition of the need for approaches that will lead to improved consistency of the teachers' judgements against the standards.

Closely connected with this finding is the current interest in the multiple purposes of assessment, and the role of assessment in improving learning. The links between assessment and learning are widely described in terms of assessment for learning, assessment of learning, and assessment as learning. The work of the *Assessment Reform Group* in the UK has been widely disseminated, and the emphasis on assessment for learning has been increasing. We found frequent reference to this work in documentation made available to teachers by educational jurisdictions. Evidence from research (Black & Wiliam, 1998a; OECD, 2005) clearly indicates that improving formative assessment improves learning.

Evidence from the rigorous systematic review conducted by the EPPI-Centre at the University of London emphasises the importance of providing professional development for teachers in undertaking assessment for different purposes, and that the process of moderation should be seen as an important means of developing teachers' understanding of learning goals and related assessment criteria. This review also found that there should be wider recognition that clarity about learning goals is needed for dependable assessment by teachers.

In Victoria, a coherent approach linking to and building upon the Victorian Essential Learning Standards (VELS), progression points, continua, assessment maps and other policy, advice and support materials has been developed and implemented. This work is of high quality, and draws from the body of research and evidence identified in the investigation. For example, the characteristics of effective assessment detailed in the Department of Education's Assessment Advice make the best of recent research available to teachers. The examples of good assessment practice collected from schools and shared both through VCAA support materials and the Department of Education's Knowledge Bank also exemplify the soundness of approach. However,

there is no systematic professional learning program related to teacher assessment in place in Victoria, nor any systematic state-wide moderation processes.

The investigation found strong links between best-practice approaches to building and improving the consistency of teacher judgements against state-wide standards and a strong focus on teacher professional learning. It was also found that teachers' participation in activities intended to improve the consistency of teacher judgements, such as moderation meetings, also leads to significant professional learning. There is a clear consensus that productive and sustainable changes in teachers' knowledge and practice stem from professional activities intended to enhance the consistency of teacher judgements. This is an important additional benefit accruing from work on making consistent judgements.

The implementation of standards-referenced curriculum planning and assessment approaches involves significant change in schools, and in teachers' practices. Such change needs time to become embedded in ongoing practice, and requires ongoing and informed leadership and support. Effective professional learning for improving the consistency of teacher judgement needs to include several components.

The Australian and international experience indicates the need to take a comprehensive approach to improving the consistency of teachers' judgements. No single element, no matter how well designed, will be sufficient on its own. The following recommendations are framed with a view to building up a high-quality framework in Victoria.

Two recommendations have been derived from the evidence of effective approaches to teacher assessment. The first identifies the key features of the best practices that should be adopted in Victoria. The second recommendation suggests the scope of initiatives required to put the first recommendation into practice.

Recommendation 1

It is recommended that the key features of best-practice approaches to improving consistency of teacher judgements against statewide standards identified in the investigation be adopted in the Victorian context:

- a. processes for developing teachers' understanding of the standards*
- b. processes for assisting teachers to identify appropriate evidence of students' achievement of the standards*
- c. processes and protocols for teachers to meet together to examine samples of their own students' work against samples demonstrating achievement at the VELS standards and progression points, and to moderate their assessments of their own students' work*
- d. opportunities for moderation within schools and with other schools*
- e. leadership and support from trained assessment leaders available to all schools*
- f. processes for external validation of the accuracy of teachers' judgements after moderation*

Recommendation 2

It is recommended that the Department, in the light of the findings of the investigation, build capacity in the Victorian teaching profession in making consistent judgements by

establishing a 3-year statewide initiative funded at an appropriate level to include the following elements:

- h. The appointment of three trained assessment support officers in each region, whose role would be to provide professional development and direct support to schools.*
- i. The provision, through the statewide/regional network of 3 days of professional development for all teachers in Years 1, 3, 5, 7 and 9 in 2008, and 3 days for all teachers in Years 2, 4, 6, 8 and 10 in 2009*
- j. The inclusion in the professional development days of activities designed to develop common understandings of the VELS, and on how to collect an appropriate body of evidence on which to base assessments*
- k. The inclusion in the professional development days of opportunities for all participating teachers to engage in moderation activities across schools.*
- l. Allocation of funding for an additional 2 student free days for all schools (matching the targeted year levels) for school level moderation, supported by regional officers.*
- m. The review, by the teams of regional assessment support officers of a random sample of assessments from approximately 20% of schools in each region in order to monitor consistency and accuracy and to validate teachers' assessments*
- n. The establishment, at the inception of the statewide assessment training initiative, of a well-defined process for reviewing the outcomes of the initiative.*

One of the major purposes of such a program of professional learning should be the development of individual teachers' knowledge and understanding of the standards, and the development of common understandings of the standards within schools, and across schools and regions. Equally important will be the development of teachers' understanding of how to plan learning opportunities that will enable students to develop the knowledge and skills defined in the standards. Further, teachers need to develop knowledge about the range of student work that can be collected as evidence of achieving a standard, or of having reached a certain point in progressing towards achieving a standard. For example, draft and finished work, print, non-print and oral work samples, and print or digital portfolios all count as 'bodies of evidence'.

A wide range of professional learning activities, whether face-to-face or on-line, will support the achievement of these purposes. These activities should include opportunities for the shared exploration of annotated work samples. These work samples should demonstrate the full range of achievement of the standards, not only samples derived from the relatively narrow confines of what can be tested under standard conditions. Teachers also need access to opportunities to collaborate in the development of assessment tasks, and to engage in assessing the work of their own students and other teachers' students against the standards. The introduction of protocols and processes for effective moderation practices within and across schools will also be significant.

This professional learning will take place in a range of contexts: school-level, cluster/network, regional and state-wide. The nature of the activity will determine the most appropriate context for particular professional learning activities

The findings from the investigation identify a number of approaches have been identified as being worthy of further consideration. They provide some signposts to the further development of effective approaches for improving the consistency of teacher judgement. The NZ *asTTLe* is an excellent example of a research-based tool developed for classroom use that has multiple

possibilities for improving learning. The identified approaches encompass teacher professional learning, the strengthening of classroom assessment practices, the effective use of planning and assessment resources, the development of strategies for collecting evidence of improvement in making consistent judgements, and in the implementation of moderation processes within and across schools.

The development of consistency in teacher judgements involves a strengthening of classroom assessment practices in schools, in particular formative assessment practices. The investigation found that a shift of emphasis to classroom assessment was prevalent in most jurisdictions.

The availability of assessment advice and a range of key resources was identified in the investigation of the current Victorian context, and the relevance of these resources to the enhancement of assessment practices is clearly apparent. The development and provision of these resources is indicative of the current priority given to effective assessment practices.

Recommendation 3

It is recommended that the strengthening of assessment practices and making consistent judgements be included as key objectives within other system initiatives, for example, the performance and development culture in Victorian schools, and the improvement of pre-service teacher education programs.

The consistency of the judgements teachers make against a set of standards lies in the extent to which the judgements are consistent over time, and the judgements of teachers in different schools are consistent. This involves:

- judgements made for individual students are consistent with the judgements previously made for those students;
- judgements for all students in a range of contexts are consistent; and
- judgements made by teachers are consistent with judgements made for all students.

Evidence that can be interrogated to determine the extent of consistency is difficult to obtain. The available evidence in Victoria lies in several years of teachers' reported judgements against CSF1 and 11, and the VELS. This reported information may show patterns, such as a similarity in the reported levels of achievement for students in a particular year level at one school over several years. It can be compared to other data for the same group of students, such as AIM results, and discrepancies between the reported achievement shown in the teacher judgements and the achievement in AIM can be identified. However, this evidence does not enable scrutiny of the aspects of consistency discussed above.

The continuation of the current practice of comparing judgements of teachers' judgement against the VELS with AIM data will provide useful information to school leaders, and signal when there may be a need to refine teachers' judgements against the standards. In addition, some longitudinal studies of the judgements made by all teachers in selected schools, or all teachers at the same year level in several schools in different metropolitan and non-metropolitan locations, would help to generate further evidence of improvements in the consistency of teacher judgements. A longitudinal approach is the only appropriate way of gathering evidence of the consistency of judgements over time.

Recommendation 4

It is recommended that current processes involving the comparison of aggregated VELS data with AIM data should continue, and that longitudinal evidence of the consistency and

accuracy of teachers' judgements be collected from an intensive study of a statewide sample of individual schools and individual teachers.

It is recommended that current processes for comparing aggregated VELS data with AIM data continue, and that longitudinal evidence be collected from an intensive study of a sample of individual schools and individual schools.

The investigation highlighted the significance of moderation processes in developing consistency of teacher judgement. It also identified the quality assurance function of moderation of teachers' judgements. Although there are resources to support moderation processes, there is currently no systematic process of moderation in place in Victoria. The investigation identified examples of effective cluster and school-based moderation processes, including the systematic large-scale moderation program across all schools as was found in Tasmania and Western Australia. The QMAP work in Tasmania in 2005-2006, for example, involved all schools working in clusters, provided extensive support in the form of assessment support leaders, and a range of resources, including *Guiding Assessment Tasks*, and time for all teachers to participate in school, cluster and state-wide moderation days. Schools in South Australia can access, on request, a moderation process provided by a central team. An interesting feature of these moderation processes was the use of moderation protocols designed to foster practices in which teachers are respectful of each others' work, and prepared to modify their own judgements through shared discussions of student work.

The literature indicates that teachers' participation in moderation activities within schools and across schools is a key element in developing consistency of teacher judgements. Moderation is known as an effective means of improving the quality of assessment in schools and of teachers' judgements, and offers significant benefits to the system as a whole. The possibilities of incorporating this key element into existing Victorian structures and initiatives could be explored. Monitoring of a range of current initiatives in a variety of contexts would identify opportunities within these initiatives in which opportunities for professional learning about assessment for learning and moderation for consistency could be integrated in the initiative. For example, the two-year DoE *Literacy Improvement Teams Initiative* has an extensive professional learning program for the Literacy Specialists who will work in schools, and a focus on assessment and moderation would fit well. Access to a small team of centrally-trained, regionally-based moderation officers to provide advice and leadership to schools would support an organised and systematic approach to developing moderation processes as part of a broader assessment culture. The provision of time for teachers to participate in moderation meetings would be a major cost, but the benefits to the system in terms of strengthening assessment practices and enhancing teachers' professional knowledge and practice would be significant.

The evidence from research clearly indicates the significance of moderation activities that involve teachers in the collaborative examination of students' work. This form of moderation is sometimes described as 'social moderation' by contrast with other forms of moderation such as statistical moderation, and moderation by visitation.

Moderation processes involving teachers' collaborative examination of samples of student work should become a central aspect of schools' assessment culture and practices. Opportunities for teachers to engage in moderation activities within and between schools, in face-to-face and online forms should be implemented as a system-wide initiative. This needs to be complemented by an external means of monitoring teachers' assessments over time.

Recommendation 5

It is recommended that the relevance of assessment policies and practices be linked to the developing performance and development culture in Victorian schools.

Recommendation 6

It is recommended that a systematic program of moderation activities involving teachers meeting together to assess samples of student work be adopted for implementation within and across Victorian schools.

The development of a comprehensive strategy needs to be informed by the best available evidence on the benefits and costs of alternative approaches. A systematic evaluation strategy needs to be embedded in the framework from the outset. This would include the use of pilot studies to assess the extent to which initiatives improve teacher practice, the conditions that lead to beneficial outcomes, and the resources involved and their costs.

Recommendation 7

It is recommended that a comprehensive strategy for improving the consistency of teachers' judgements in Victorian schools be developed, and that this strategy should be informed and supported by a systematic evaluation of impact, benefits and costs.

REFERENCES

- Allen, J. (2003). Personal communication with the Director, Assessment and New Basics Branch. Cited in the *New Basics Research Report*. Brisbane: Queensland Government 2004.
- Assessment and New Basics Branch. (2004). *New Basics Research Report*. Retrieved. from.
- Black, P. (1995). Curriculum and Assessment in Science Education: The Policy Interface. *International Journal of Science Education*, 17(4), 453-469.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for Learning: Putting it into practice*. Maidenhead: Open University Press.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 86(1), 9-21.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Bruniges, M. (2001). *The consistency and comparability of teacher judgement*. Paper presented at the Consistency of Outcomes in Stage Assessment.
- Condie, R., Livingstone, K., & Seagraves, L. (2005). *Evaluation of the Assessment is for Learning Programme: Final Report and Appendices*. Retrieved. from.
- Consistency Project Report. *Enhancing Consistency in Teacher Judgement of Student Work. Victorian report*. (1998). Retrieved. from.
- Council of Ontario Directors of Education. (2006). *Consistency in Classroom Assessment: Support Materials for Educators*. Retrieved. from.
- Council of Ontario Directors of Education. (May 2006). *Consistency in Classroom Assessment: Support Materials for Educators*. Retrieved. from.
- Cumming, J. J., & Maxwell, G. S. (2004). Profiles of educational assessment systems worldwide. *Assessment in Education: principles, policy & practice*, 11(1), 89-108.
- Earl, L. M. (2003). *Assessment as Learning*. Thousand Oaks: Corwin Press.
- Gipps, C., Clarke, S., & McCallum, B. (1998). *The Role of Teachers in National Assessment in England*. Paper presented at the Annual Meeting of the American Educational Research Association.
- Harlen, W. (1994). *Concepts of Quality in Student Assessment*. Paper presented at the Annual Meeting of the American Educational Research Association. from
- Harlen, W. (2004). *A systematic review of the evidence of reliability and validity of assessment by teachers used for summative purposes*. London: EPPI-Centre, Social Science Research Unit, Institute of Education.
- Harlen, W. (2005a). Teachers' summative practices and assessment for learning - tensions and synergies. *Curriculum Journal*, 16(2), 207-223.
- Harlen, W. (2005b). Trusting teachers' judgement: research evidence of the reliability and validity of teachers' assessment used for summative purposes. *Research Papers in Education*, 20(3), 245-270.
- Hattie, J., & Jaeger, R. (1998). Assessment and classroom learning: a deductive approach. *Assessment in Education: principles, policy & practice*, 5(1), 111-122.
- Koh, K., Lee, A. N., Gong, W., & Wong, H. M. (2006). *Development of the Singapore Prototype Classroom Assessment Tasks: Innovative Tools for Improving Student Learning and Performance*. Paper presented at the International Association for Educational Assessment 32nd Annual Conference.
- Little, J. W., Gearhart, M., Curry, M., & Kafka, J. (2003). Looking at Student Work For Teacher Learning, Teacher Community, and School Reform. *Phi Delta Kappan*, 85(3), 184-192.
- Loucks-Horsley, S., Hewson, P. W., Love, N., & Stiles, K. E. (1998). *Designing professional development for teachers of science and mathematics*. Thousand Oaks: Corwin Press.
- Marshall, B. (2004). Goals or horizons--the conundrum of progression in English: or a possible way of understanding formative assessment in English. *Curriculum Journal*, 15(2), 101-113.
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with Assessment for Learning: lessons from the classroom. *Research Papers in Education*, 21(2), 133-149.

- Matters, G. (2004). *Two out of five is bad*. Paper presented at the Anglican Schools Conference, Brisbane.
- Matters, G. (2005). *Good data, bad news, and good policy making*. Paper presented at the ACER Research Conference 2005: Using data to support learning.
- Matters, G. (2006). *Using Data to Support Learning in Schools: Students, teachers, systems*. Camberwell: ACER Press.
- Maxwell, G. S. (2002). *Moderation of teacher judgements in student assessment*. Retrieved from National Assessment Agency. (2004). *Building a picture of what children can do*. London: National Assessment Agency.
- OECD. (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris: OECD.
- Qualifications and Curriculum Authority. (2006). *Assessment and reporting arrangements: Key stage 1: 2007*. Qualification and Curriculum Authority.
- Reed, M., & Lewis, K. (2005). *Key Stage 1 Evaluation of New Assessment Arrangements*: National Assessment Agency.
- Ryan, T. G. (2006). Performance Assessment: Critics, Criticism and Controversy. *International Journal of Testing*, 6(1), 97-104.
- Sadler, D. R. (1987). Specifying and Promulgating Achievement Standards. *Oxford Review of Education*, 13(2), 191-209.
- Taggart, G., Whitby, K., & Sharp, C. (2004). *Curriculum and Progression in the Arts: an International Study: Final Report*: National Foundation for Educational Research.
- Tierney, R. D. (2006). Changing practices: influences on classroom assessment. *Assessment in Education: principles, policy & practice*, 13(3), 239-264.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, 70(9), 703-713.

ADDITIONAL RESEARCH AND POLICY DOCUMENTS

- Allen, D., & McDonald, J. (2003, 21 January 2003). The tuning protocol: A process for reflection on teacher and student work. Retrieved November 3, 2006, from http://www.essentialschools.org/cs/resources/view/ces_res/54
- Arkoudis, S., & O'Loughlin, K. (2004). Tensions between validity and outcomes: Teacher assessment of written work of recently arrived immigrant ESL students. *Language Testing*, 21(3), 284 - 304.
- Assessment Reform Group. (1999). *Assessment for learning: Beyond the black box*. Cambridge: Cambridge University Faculty of Education.
- Assessment Reform Group. (2002). *Testing, motivation and learning*. Cambridge: University of Cambridge Faculty of Education.
- Association of Assessment Inspectors and Advisers. Teacher assessment in action. Retrieved November 3, 2006, from <http://www.aaia.org.uk/aaiaainaction.html>
- Bennett, J., & Taylor, C. (2003). Is assessment for learning in a high-stakes environment a reasonable expectation? Australasian Curriculum, Assessment and Certification Authorities National Conference. Adelaide, South Australia.
- Black, P., & Wiliam, D. (2005). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices. *Curriculum Journal*, 16(2), 249 – 261.
- Buckendahl, C. W., Impara, J. C., & Plake, B. S. (2001). Strategy for evaluating district developed assessments for state accountability, Midwestern Educational Research Association Annual Meeting. Chicago.
- Budge, D. (2006, 2/3/2006). Judgements under fire. *Times Educational Supplement*, p. 13.
- Castleton, G., Wyatt-Smith, C., Cooksey, R., & Freebody, P. (2003). The nature of teachers' qualitative judgements: A matter of context and salience: Part two: Out-of-context judgements. *Australian Journal of Language and Literacy*, 26(2), 33-42.
- Cushing, C. (2002). A reflection on developing a consistency of teacher judgement process. *Curriculum Matters*, 1(3), 10.
- Cushman, K. (1996). Looking collaboratively at student work: An essential toolkit. *Horace*, 13(2). Department of Education Tasmania. (25 August 2005). Professional learning: Protocols. Retrieved November 3, 2006, from <http://www.ltag.education.tas.gov.au/Proflearn/commlearn/protocols.htm>
- Dompnier, B., Pansu, P., & Bressoux, P. (2006). An integrative model of scholastic judgements: Pupils' characteristics, class context, halo effect and internal attributions. *European Journal of Psychology of Education*, 21(2), 119-133.
- Forster, M. (2006). Assessment and reporting: What do we need to understand? *Professional Voice*, 4(2), 21-25.
- Freebody, P. P. (2005). *Background, rationale and specifications: Queensland curriculum, assessment and reporting framework*: Department of Education and the Arts, Queensland Government.
- Garcia, J., & Rothman, R. (2002). *Three paths, one destination: Standards-based reform in Maryland, Massachusetts and Texas*. Washington DC: Achieve Inc.
- Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment*. London: The Falmer Press.
- Gipps, C., & Clarke, S. (1998). *Monitoring consistency in teacher assessment and the impact of SCAA 's guidance materials at key stages 1, 2, and 3: Final report*. London: QCA.
- Guskey, T. R. (2001). Helping standards make the grade. *Educational Leadership*, 59(1), 20-27.
- Hall, K., & Harding, A. (2002). Level descriptions and teacher assessment in England: Towards a community of assessment practice. *Educational Research*, 44(1), 1-15.

- Hall, K., Webber, B., Varley, S., Young, V., & Dorman, P. (1997). A study of teacher assessment at key stage 1. *Cambridge Journal of Education*, 27(1), 107–122.
- Hargreaves, A., Moore, S., & James-Wilson, S. (1997). *How teachers deal with secondary school change*. Toronto: Ontario Institute for Studies in Education of the University of Toronto.
- Harrison, C., Black, P.J., Hodgen, J., & Serret, N. (2007) Strengthening Teacher Assessment Practices. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, April 9-13 2007
- Hattie, J., & Jaeger, R. (1998). Assessment and classroom learning: A deductive approach. *Assessment in Education: principles, policy & practice*, 5(1), 111-122.
- Hofkins, D. (2004). Is SAT-free the answer? *Times Educational Supplement* (4584), 22.
- James, M., Black, P., McCormick, R., Pedder, D., & Wiliam, D. (2006). Learning how to learn, in classrooms, schools and networks: Aims, design and analysis. *Research Papers in Education*, 21(2), 101-118.
- King, D. (2005). *Position paper: Consistency of teacher judgement for years 1 to 10 in Lutheran schools: Sharing understandings - developing professionalism*: Lutheran Education Queensland.
- Linn, R. L. (1993). Linking results of distinct assessments. *Applied Measurement in Education*, 6(1), 83-102.
- Linn, R. L., & Baker, E. L. (1993). Comparing results from disparate assessments. *The CRESST Line*, Winter, 1-2.
- Lowery, V. (2006). *Consistency of teacher judgement - it's a bit tricky in a small school*: NSW Department of Education and Training.
- Morgan, C., & Watson, A. (2002). The interpretive nature of teachers' assessment of students' mathematics: Issues for equity. *Journal for Research in Mathematics Education*, 33(2), 78.
- Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research*, 62(3), 229-258.
- Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, 23(2), 5-12.
- New South Wales Department of Education and Training. (2003). Consistent teacher judgement online. Retrieved November 2, 2006, from <http://10ss.qtp.nsw.edu.au/ctj/>
- Office for Standards in Education. (2005). *English 2000-05: A review of inspection evidence*: Office for Standards in Education.
- Office of Learning and Teaching. (2006). Assessment and reporting - student learning - blueprint. Retrieved November, 2006, from <http://www.sofweb.vic.edu.au/blueprint/fs1/assessment.asp>
- Office of the Queensland School Curriculum Council. (2002). *Consistency of teacher judgment research report*. Brisbane: Queensland School Curriculum Council.
- Painter, J. (2006). End of key stage 1 assessment: Teacher assessment and moderation 2006. Retrieved December, 2006, from <http://education.staffordshire.gov.uk/NR/rdonlyres/2976A61B-A312-4C6D-9106-CDC6CD0FD56A/31302/Finalisingteacherassessmentjudgementsandmoderation.doc>
- Pentti, A. (2004). Assessment in the context of moderation. *Curriculum Matters*, 3(2), 18-19.
- Perry, N. E., & Meisels, S. J. (1996). *How accurate are teacher judgements of students' academic performance?* (No. 96-08 Working Paper Series). Washington: National Center for Education Statistics.
- Poskitt, J., Antony, G., Brown, M., & Taylor, K. (2002). *Research on the development of national exemplars: Implications for teacher education in New Zealand*, Teacher Education Forum of Aotearoa New Zealand Conference. Wellington, New Zealand.
- Professional Support and Curriculum Directorate. (2002). *Consistent teacher judgment in action: A resource for schools*: New South Wales Department of Education and Training.

- Queensland School Curriculum Council. (2002). *Consistency of teacher judgment*. Brisbane: Queensland School Curriculum Council.
- SACSA Website Team. Accountability and assessment: Consistency of teacher judgment. Retrieved December, 2006, from <http://www.sacsa.sa.edu.au/link.asp?ID=FRAMEWORK:G3A:TAG>
- Sadler, D. R. (1989). Formative assessment in the design of instructional systems. *Instructional Science*, 18, 119-144.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: principles, policy & practice*, 5(1), 77-84.
- Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Smith, C. B. E. (2003). *Alternative forms of assessment*. ERIC topical bibliography and commentary. Bloomington: ERIC Clearinghouse on Reading, English and Communication.
- South Australian Department of Education and Children's Services. (2001a). Accountability and assessment PD modules. Retrieved November 2, 2006, from http://www.sacsa.sa.edu.au/index_fsrc.asp?t=AA
- South Australian Department of Education and Children's Services. (2001b). Consistency of teacher judgment. Retrieved October 24, 2006, from http://www.sacsa.sa.edu.au/index_fsrc.asp?t=AA
- South Australian Department of Education and Children's Services. (2006a). *Let's Talk Assessment*, 1(3).
- South Australian Department of Education and Children's Services. (2006b). *Let's Talk Assessment*, 1(2).
- South Australian Department of Education and Children's Services. (2006c). *Let's Talk Assessment*, 1(1).
- South Australian Department of Education Training and Employment. (2000). Consistency of teacher judgement - a training and development CD-ROM for teachers. South Australia: MindVision Interactive.
- Stables, A. (1992). Speaking and listening at key stage 3: Some problems of teacher assessment. *Educational Research*, 34, 107-115.
- Stokking, K., van der Schaaf, M., Jaspers, J., & Erkens, G. (2004). Teachers' assessments of students' research skills. *British Educational Research Journal*, 30(1), 93-166.
- Thomas, G., Ward, J., & Tagg, A. (2005). Numeracy assessment: How reliable are teachers' judgments? In *Findings from the New Zealand numeracy development projects 2005*. Wellington: Learning Media Limited for the Ministry of Education.
- Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal*, 22(4), 389-404.
- Valencia, S. W., & Calfee, R. (1991). The development and use of portfolios for students, classes and teachers. *Applied Measurement in Education*, 4(4), 333-345.
- Webb, M. & Jones, J. (2007) Developing Assessment for learning: How teachers' and students' beliefs and experiences affect changing classroom practice. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, April 9-13 2007
- Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey Bass.
- Wiliam, D. (2003). National curriculum assessment: How to make it better. *Research Papers in Education*, 18(2), 129-136.

- Wilson, M. (2004). Community of judgment: A teacher-centered approach to educational accountability. In Office of Technology Assessment (Ed.), *Issues in educational accountability*. Washington DC: Office of Technology Assessment, United States Congress.
- Wyatt-Smith, C. (2005). Being serious about improving student outcomes: Bringing the issues of quality, standards, and teacher judgement to centre stage. In B. Bartlett, F. Bryer & D. Roebuck (Eds.), *Stimulating the "Action" as Participants in Participatory Research: International Conference on Cognition, Language and Special Education Research* (Vol. 1, pp. 25-38). Surfers Paradise: Griffith University, School of Cognition, Language, and Special Education.
- Wyatt-Smith, C., & Castleton, G. (2005). Examining how teachers judge student writing: An Australian case study. *Journal of Curriculum Studies*, 37(2), 131-154.
- Wyatt-Smith, C., Castleton, G., Freebody, P., & Cooksey, R. (2003). The nature of teachers' qualitative judgements: A matter of context and salience: Part one: 'in-context' judgement. *Australian Journal of Language and Literacy*, 26(2), 11-32.
- Zhicheng, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education*, 16(4), 323-342.

Government and Curriculum Documents

- ACT Department of Education and Training. (2006). *Every chance to learn: Curriculum framework for ACT schools: Preschool to year 10: Phase 2 - trial draft*: Department of Education and Training.
- Board of Studies New South Wales. (2005). *NSW primary curriculum foundation statements*: Department of Education and Training.
- Council of Ontario Directors of Education. (2006). *Consistency in classroom assessment: Support materials for educators*: Ministry of Education.
- Department for Education and Skills. (2003). *Excellence and enjoyment: A strategy for primary schools*. London: DfES.
- Department for Education Training and Employment South Australia. (2001). *South Australian curriculum, standards and accountability framework*. Department for Education Training and Employment.
- Department of Education and Training Western Australia. *Outcomes and standards framework*: Department of Education and Training.
- Department of Education and Training Western Australia. (2005). *Annual report 2004/2005*: Department of Education and Training.
- Department of Education and Training Western Australia. (2006). *Curriculum, assessment and reporting K-10: Policy and guidelines*: Department of Education and Training.
- Department of Education Queensland. (2002). *The report of the assessment and reporting taskforce*: The State of Queensland (Department of Education).
- Department of Education Tasmania. *Essential Learnings Framework*: Department of Education.
- Department of Education Tasmania. (2002). *Assessing and reporting student achievement against the Essential Learnings Framework: Assessment, monitoring and reporting policy*: Department of Education Tasmania.
- Department of Education Tasmania. (2005). *Essential learnings assessing guide*.
- Department of Education Tasmania. (2006). *Annual report 2005-6*: Department of Education.
- Department of Employment Education and Training Northern Territory. *Northern Territory curriculum framework*: Department of Employment Education and Training.
- Nelson, J. (2006). Key stage 1 statutory assessment 2005-6: Medway guide: Medway Council.

- New South Wales Department of Education and Training. (2005). *Getting the balance right: The department's response to time to teach, time to learn*: New South Wales Department of Education and Training.
- New Zealand curriculum framework*. (1993). New Zealand Ministry of Education.
- The New Zealand curriculum: Draft 2006–2007*. (2006). New Zealand Ministry of Education.
- Office of Learning and Teaching. (2006). *Closing the loop: Curriculum, pedagogy, assessment & reporting*: Victorian Department of Education & Training.
- Ontario Ministry of Education. *The Ontario curriculum*: Ministry of Education.
- Ontario Ministry of Education. (2000). *The Ontario curriculum, grades 9 to 12: Program planning and assessment*: Ministry of Education.
- Ontario Ministry of Education. (2002). *The Ontario curriculum unit planner: Assessment strategies companion*: Queens Printer for Ontario.
- Qualifications and Curriculum Authority. *The national curriculum*: Department for Education and Employment.
- Qualifications and Curriculum Authority. (2006). *Assessment and reporting arrangements: Key stage 1: 2007*: Qualification and Curriculum Authority.
- Queensland Department of Education and the Arts. (2005). *Queensland curriculum, assessment and reporting framework: Strategic Policy and Education Futures*, Department of Education and the Arts.
- South Australian Department of Education and Children's Services. *SACSA moderation protocol*: South Australian Department of Education and Children's Services.
- Victorian Curriculum and Assessment Authority. *Victorian Essential Learning Standards*: Department of Education.