Unclassified **EDU/EDPC/EA(2011)10**

Organisation de Coopération et de Développement Économiques
Organisation for Economic Co-operation and Development **06-Dec-2011**

English text only

DIRECTORATE FOR EDUCATION
EDUCATION POLICY COMMITTEE

EDU/EDPC/EA(2011)10
Unclassified

Cancels & replaces the same document of 18 August 2011

Group of National Experts on Evaluation and Assessment

OECD Reviews of Evaluation and Assessment in Education: Australia

by Paulo Santiago, Graham Donaldson, Joan Herman and Claire Shewbridge

The Review of Evaluation and Assessment in Education: Australia forms part of the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes. This report is only available in PDF format and can also be downloaded from the project's website at www.oecd.org/edu/evaluationpolicy.

Contacts: Mr. Paulo Santiago [Tel: +33 (0)1 45 24 84 19; e-mail: paulo.santiago@oecd.org];
Ms. Claire Shewbridge [Tel: +33 (0)1 45 24 99 63; e-mail: claire.shewbridge@oecd.org].

JT03312956

English text only

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

OECD Reviews of Evaluation and Assessment in Education

AUSTRALIA

Paulo Santiago, Graham Donaldson,
Joan Herman and Claire Shewbridge

OECD

OECD Reviews of Evaluation and Assessment in Education

Australia

Paulo Santiago, Graham Donaldson,
Joan Herman and Claire Shewbridge

August 2011

OECD

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 34 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and the arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

© OECD 2011

No translation of this document may be made without written permission. Applications should be sent to **rights@oecd.org**.

Foreword

This report for Australia forms part of the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes (see Annex A for further details). The purpose of the Review is to explore how systems of evaluation and assessment can be used to improve the quality, equity and efficiency of school education. The Review looks at the various components of assessment and evaluation frameworks that countries use with the objective of improving student outcomes. These include student assessment, teacher appraisal, school evaluation and system evaluation.

Australia was one of the countries which opted to participate in the country review strand and host a visit by an external review team. Members of the Review Team were Paulo Santiago (OECD Secretariat), co-ordinator of the Review; Graham Donaldson (formerly Her Majesty's Senior Chief Inspector of Education in Scotland; United Kingdom); Joan Herman (Director, National Center for Research on Evaluation, Standards, and Student Testing at the University of California – Los Angeles, United States); and Claire Shewbridge (OECD Secretariat). This publication is the report from the Review Team. It provides, from an international perspective, an independent analysis of major issues facing the evaluation and assessment framework in Australia, current policy initiatives, and possible future approaches. The report serves three purposes: (1) Provide insights and advice to the Australian education authorities; (2) Help other OECD countries understand the Australian approach; and (3) Provide input for the final comparative report of the project.

Australia's involvement in the OECD Review was co-ordinated by Ms. Kristie van Omme, School Improvement and Transparency Branch, Curriculum, Assessment and Teaching Group, Australian Government Department of Education, Employment and Workplace Relations (DEEWR). An important part of Australia's involvement was the preparation of a comprehensive and informative Country Background Report (CBR) on evaluation and assessment policy, published by the Australian Government in 2010. The Review Team is very grateful to the authors of the CBR, and to all those who assisted them for providing an informative document. The CBR is an important output from the OECD Review in its own right as well as an important source for the Review Team. Unless indicated otherwise, the data for this report are taken from the Australian Country Background Report. The CBR follows guidelines prepared by the OECD Secretariat and provides extensive information, analysis and discussion in regard to the national context, the organisation of the educational system, the main features of the evaluation and assessment framework and the views of key stakeholders. In this sense, the CBR and this report complement each other and, for a more comprehensive view of evaluation and assessment in Australia, should be read in conjunction.

The Review Visit to Australia took place on 21-30 June 2010 and covered visits to Canberra and the states of New South Wales, Queensland, Victoria and Western Australia. The itinerary is provided in Annex B. The visit was designed by the OECD in

collaboration with the Australian authorities. The biographies of the members of the Review Team are provided in Annex C.

During the review visit, the team held discussions with a wide range of national, state and territory authorities; statutory bodies at the national and systemic levels; teacher unions; parents' organisations; representatives of principals; students; representatives of employers; and researchers with an interest in evaluation and assessment issues. The team also visited a range of schools, interacting with school management, teachers and students. The intention was to provide a broad cross-section of information and opinions on evaluation and assessment policies and how their effectiveness can be improved.

The Review Team wishes to record its grateful appreciation to the many people who gave time from their busy schedules to assist in its work. The education community clearly attached great importance to the purpose of the visit and the fact that the Review Team brought an external perspective. The meetings were open and provided a wealth of information and analysis. Special words of appreciation are due to the National Co-ordinator, Kristie van Omme, and her colleague Jessica Yelavich, Curriculum, Assessment and Teaching Group, DEEWR, for going to great lengths to respond to the questions and needs of the Review Team. We were impressed by their efficiency and expertise and enjoyed their kindness and very pleasant company. The courtesy and hospitality extended to us throughout our stay in Australia made our task as a Review Team as pleasant and enjoyable as it was stimulating and challenging. The OECD Review Team is also grateful to colleagues at the OECD, especially to Stefanie Dufaux for preparing the statistical annex to this report (Annex D) and to Heike-Daniela Herzog for editorial support.

This report is organised in six chapters. Chapter 1 provides the national context, with information on the Australian school system and the main recent developments. Chapter 2 looks at the overall evaluation and assessment framework and analyses how the different components of the framework play together and can be made more coherent to effectively improve student learning. Then Chapters 3 to 6 present each of the key components of the evaluation and assessment framework – student assessment, teacher appraisal, school evaluation and system evaluation – in more depth, presenting strengths, challenges and policy recommendations.

The policy recommendations attempt to build on and strengthen reforms that are already underway in Australia, and the strong commitment to further improvement that was evident among those we met. The suggestions should take into account the difficulties that face any visiting group, no matter how well briefed, in grasping the complexity of Australia and fully understanding all the issues.

Of course, this report is the responsibility of the Review Team. While we benefited greatly from the Australian CBR and other documents, as well as the many discussions with a wide range of Australian personnel, any errors or misinterpretations in this report are our responsibility.

Table of contents

Acronyms and abbreviations **7**

Executive summary **9**

Chapter 1: School education in Australia **13**

Main features of the school system 14
Main policy developments: the national agenda for education 18
References 23

Chapter 2: The evaluation and assessment framework **25**

Context and features 26
Strengths 30
Challenges 36
Policy recommendations 40
References 46

Chapter 3: Student assessment **47**

Context and features 48
Strengths 55
Challenges 59
Policy recommendations 63
References 70

Chapter 4: Teacher appraisal **73**

Context and features 74
Strengths 83
Challenges 86
Policy recommendations 91
References 97

Chapter 5: School evaluation **99**

Context and features 100
Strengths 106
Challenges 110
Policy recommendations 114
References 120

Chapter 6: Education system evaluation **121**

Context and features 122
Strengths 125
Challenges 131
Policy recommendations 134
References 142

Conclusions and recommendations **145**

Education system context 145
Strengths and challenges 146
Policy recommendations 152

Annex A: The OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes **159**

Annex B: Visit itinerary **161**

Annex C: Composition of the Review Team **171**

Annex D: Comparative indicators on evaluation and assessment **173**

Tables

Table 1.1 Structure of school education across states and territories 14
Table 1.2 COAG's National Productivity Agenda – Schools 19
Table 2.1 Levels of evaluation and assessment in Australia 29
Table 4.1 Levels of teacher registration/accreditation across Australia 75
Table 4.2 Framework for teachers' performance management in government schools across Australian jurisdictions 77
Table 4.3 Advanced Skills Teaching schemes in government schools across Australian jurisdictions 79

Boxes

Box 2.1 Equity within the performance framework 32
Box 2.2 Data information systems in New South Wales and Victoria 35
Box 2.3 Strengthening monitoring on literacy and numeracy in Ontario 44
Box 3.1 Validity criteria to judge individual student assessments 54
Box 3.2 Strategies to improve the reliability of teacher-based summative assessment 66
Box 3.3 Individual Development Plans for students in Sweden 69
Box 4.1 The teaching profession in Australia – Main features 81
Box 4.1 The teaching profession in Australia – Main features *(continued)* 82
Box 5.1 Information reported on the *My School* website 105
Box 6.1 Sample surveys in the Netherlands and New Zealand 138
Box 6.2 Use of data systems for decision making by educational districts in the United States 139
Box 6.3 Proposed monitoring of schools in all sectors in Victoria 140

Acronyms and abbreviations

ABS	Australian Bureau of Statistics
ACACA	Australasian Curriculum, Assessment and Certification Authorities
ACARA	Australian Curriculum, Assessment and Reporting Authority
ACER	Australian Council for Educational Research
ACT	Australian Capital Territory
AITSL	Australian Institute of Teaching and School Leadership
ALGA	Australian Local Government Association
AQF	Australian Qualifications Framework
AST	Advanced Skills Teaching
ATAR	Australian Tertiary Admission Rank
BARS	Behaviourally Anchored Rating Scales
CBR	Country Background Report
COAG	Council of Australian Governments
DEEWR	Department of Education, Employment and Workplace Relations (Australian Government)
EOL	English Online Interview
ESSA	Essential Secondary Science Assessment
ICSEA	Index of Community Socio-Educational Advantage
ICT	Information and Communication Technologies
IEA	International Association for the Evaluation of Educational Achievement
LBOTE	Language Background Other than English
MCEECDYA	The Ministerial Council on Education, Early Childhood Development and Youth Affairs
MCEETYA	The Ministerial Council on Education, Employment, Training and Youth Affairs
NAP	National Assessment Program
NAPLAN	National Assessment Program Literacy and Numeracy
NCTES	National Centre for Education and Training Statistics
NCVER	National Centre for Vocational Education Research
NEA	National Education Agreement
NP	National Partnership
NSSC	National Schools Statistics Collection
NSW	New South Wales
NT	Northern Territory
OECD	Organisation for Economic Co-operation and Development
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
QCAT	Queensland Comparable Assessment Tasks
RTO	Registered Training Organisation
SA	South Australia
SES	Socio-economic status
SMART	School Measurement, Assessment and Reporting Toolkit
TALIS	Teaching and Learning International Survey
TIMSS	Trends in International Mathematics and Science Study
VET	Vocational Education and Training
VRQA	Victorian Registration and Qualification Authority
WA	Western Australia
WAMSE	The Western Australian Monitoring Standards in Education

Executive summary

In 2008 a major national agenda was established with a common framework for reform in education agreed between the Australian Government and the state and territory governments through the National Education Agreement (NEA). The clear and widely supported national education goals, articulated in the NEA and the Melbourne Declaration on Educational Goals for Young Australians, provide a solid reference point on which to build evaluation and assessment strategies to achieve accountability and improvement in student learning. The Australian approach combines the development of goals, monitoring and reporting at the national level with local evaluation and assessment practices shaped by jurisdiction-level school improvement frameworks. While the key elements of evaluation and assessment are well established at student, teacher, school and system levels, challenges remain in determining what constitutes a desirable measure of national consistency as against legitimate cross-jurisdiction diversity, and in articulating the different elements of the overall evaluation and assessment framework to ensure consistency and complementarity.

Establishing national strategies for strengthening the linkages to classroom practice

The overall evaluation and assessment framework appears as highly sophisticated and well conceptualised, especially at its top level (national and systemic levels). However, there is a less clear articulation of ways for the national agenda to generate improvements in classroom practice through the assessment and evaluation procedures which are closer to the place of learning. Moreover, striking the right balance between nationally-dictated policies and ability to meet local needs is a challenge and there is room to improve the integration of the non-governmental sector. Realising the full potential of the overall evaluation and assessment framework involves establishing strategies to strengthen the linkages to classroom practice. A major step in this direction would be a national reflection about the nature and purpose of evaluation components such as school evaluation, teacher appraisal and student formative assessment within the overall education reform strategy and the best approaches for these evaluation components to improve classroom practices. The agreement of protocols between educational jurisdictions and the Australian Government in these areas could also be the basis for promoting national consistency while giving room for local diversity. Requiring the non-government sector to be part of such protocols could also improve its integration in the overall evaluation and assessment framework.

Further developing articulations within the overall evaluation and assessment framework.

The process of developing an effective evaluation and assessment framework should give due attention to: achieving proper articulation between the different evaluation components (*e.g.* school evaluation and teacher appraisal); warranting the several elements within an evaluation component are sufficiently linked (*e.g.* teaching standards and teacher appraisal); and ensuring processes are in place to guarantee the consistent application of evaluation and assessment procedures (*e.g.* consistency of teachers' grades).

Maintaining the centrality of teacher-based student assessment while ensuring the diversity of assessment formats

A range of provisions for the assessment of student learning are established, which results in a coherent system that potentially can provide a comprehensive picture of student performance relative to Australia's goals for student learning. Following the introduction of the Australian Curriculum, sound strategies to assess against the standards/curriculum are paramount. The current strategy for student assessment consists of a combination of NAPLAN and teacher-based assessments against the full range of curriculum goals. The latter implies a considerable investment on teacher capacity to assess against the standards, including specific training for teachers, the development of grading criteria and the strengthening of moderation processes within and across schools. Also, the current prominence of NAPLAN within the student assessment framework requires particular care about not reducing the importance of teacher-based assessment. Another area of priority is NAPLAN's alignment with the Australian Curriculum and the extent to which NAPLAN is balanced in its representation of the depth and breadth of intended student learning goals.

Strengthening teacher appraisal

Teachers benefit from a high degree of trust and extensive autonomy, but they have few opportunities for professional feedback. Teacher appraisal as part of regular performance management processes is also of variable quality. The teaching profession would benefit from the alignment of teaching standards with a competency-based career structure for teachers. This would strengthen the incentive for teachers to improve their competencies, and reinforce the matching between teachers' levels of competence and the tasks which need to be performed in schools to improve student learning. As a result, teacher registration could be conceived as career-progression evaluation. It would have as its main purposes holding teachers accountable for their practice, determining advancement in the career, and informing the professional development plan of the teacher. Also, teacher appraisal as part of performance management processes should be conceived as developmental evaluation, *i.e.* the main process through which the improvement function of teacher appraisal is achieved. It would retain its current character but school-based processes for developmental evaluation would need to be strengthened and validated externally.

Defining the strategic purposes and scope of school evaluation

School self-evaluation is an expectation and some form of external review mechanism is increasingly common. Test results, focusing on literacy and numeracy, are widely used to inform evaluation. However, there remains a need to clarify a number of vital issues relating to the relationship between the role of reviews in both accountability and improvement, the scope of reviews in relation to the emerging national agenda, the critical areas on which reviews should focus, the role and nature of externality, and the extent of transparency. Different jurisdictions have addressed mixtures of these issues in their own context but no clear national direction of travel has as yet emerged. Moves towards achieving a much closer alignment between self-evaluation and external evaluation could prove beneficial – the central requirement is that internal evaluation and external evaluation use common criteria and share a common language of quality. The scope and frequency of external review are also important issues. The implementation of the broadening Australian Curriculum suggests a more general focus than that which a "failing schools" agenda might imply. For these reasons, developing policy on school evaluation in Australia should seek to use its potential to challenge complacency and provide evidence about progress on a broad front.

Continuing efforts to meet information needs for national monitoring and further exploiting results at systemic level

There are clear standard frameworks both for reporting key performance measures and for general government sector reporting, and a strong and stable set of national measures on education is established. Similarly, there are strong procedures for system monitoring at the state and territory level. The immediate priority for meeting information needs to adequately monitor progress towards national goals is to strengthen the information systems regarding student socio-economic and Indigenous status. In addition, states and territories should maintain efforts to strengthen monitoring structures, in part by further exploiting the analysis of results from local information systems and the national monitoring system, and importantly by ensuring adequate monitoring and follow-up on priority areas. Another area of priority should be to support and promote greater monitoring in the non-government sector.

Chapter 1

School education in Australia

Australia has a federal school system with primary responsibility for school education granted to state and territory governments. Student learning outcomes in Australia are very good by international standards even if there is evidence of some decline in the last decade. In 2008 a major national agenda was established with a common framework for reform in education agreed between the Australian Government and the state and territory governments through the National Education Agreement (NEA). It developed from the National Productivity Agenda agreed by the Council of Australian Governments and is supported by the Melbourne Declaration on Educational Goals for Young Australians, which articulates future directions and aspirations for Australian schooling. The main components of the national reform agenda are the development of the Australian Curriculum, the Aboriginal and Torres Strait Islander Education Action Plan, the National Partnerships, the National Assessment Program and the leadership of national-level entities such as the Australian Curriculum, Assessment and Reporting Authority (ACARA) and the Australian Institute of Teaching and School Leadership (AITSL). For the first time in Australia at the national level, the management of curriculum, assessment and reporting are brought together (through ACARA) and there is national leadership in the profession of teaching and school leadership (through AITSL). The NEA also brings an obligation to meet a common set of national school performance and reporting requirements. In this context, evaluation and assessment are key tools to monitor whether goals for quality and equity in education are being achieved.

Main features of the school system

Structure of school education

Australia has a federal school system which includes six states and two territories. The Constitution of Australia allocates primary responsibility for school education to state and territory governments.

Australia has both public and private schools which are usually referred to as "government" and "non-government" schools. Government schools operate under the direct responsibility of the relevant state or territory minister, while non-government schools are established and operate under conditions determined by government and state or territory registration authorities. Non-government schools can be part of a system of schools (systemic) or completely independent. Many non-government schools have some religious affiliation, most with the Catholic Church and as such, the non-government sector in Australia is often split into "Catholic" and "Independent" for reporting purposes. During 2009, 3.48 million students (including part-time students) attended school in 9 529 institutions across Australia (see Table 1 in the Country Background Report, Australian Government, 2010). Of these, 2.29 million students (66% of total) attended 6 802 government schools and 1.19 million students (34%) attended 2 727 non-government schools. Of the non-government schools, 1 705 were classified as Catholic schools and 1 022 as Independent.

The structure of school education varies across the eight states and territories. Formal school education in Australia comprises primary education and secondary education, including a pre-year 1 grade. Depending on the state or territory, primary school education consists of seven to eight years followed by five to six years of secondary school education. There are two basic patterns in current formal schooling in Australia, as reflected in Table 1.1.

Table 1.1 Structure of school education across states and territories

State or territory	Minimum school starting age	Primary education	Secondary education
Australian Capital Territory	4 years, 8 months	Kindergarten, Years 1-6	Years 7-12
New South Wales	4 years, 5 months	Kindergarten, Years 1-6	Years 7-12
Northern Territory	4 years, 6 months	Transition, Years 1-6	Years 7-12
Queensland	4 years, 6 months	Preparatory, Years 1-7	Years 8-12
South Australia	5 years	Reception, Years 1-7	Years 8-12
Tasmania	5 years	Preparatory, Years 1-6	Years 7-10 + Post-compulsory Years 11-12
Victoria	4 years, 8 months	Preparatory, Years 1-6	Years 7-12
Western Australia	4 years, 6 months	Pre-primary, Years 1-7	Years 8-12

Source: Australian Government (2010).

The Australian Qualifications Framework (AQF) outlines the hierarchy of qualifications in Australia. It is owned, supported and funded through the Ministerial Council for Tertiary Education and Employment and provides a nationally recognised

official qualification. After completing secondary schooling, most Australian students are awarded a senior secondary certificate of education, which certifies the completion of secondary education. Vocational Education and Training (VET) is also available in some schools across Australia. VET programmes undertaken by school students as part of a senior secondary certificate provide credit towards a nationally recognised VET qualification within the Australian Qualifications Framework. The training that students receive reflects specific industry competency standards and is delivered by a Registered Training Organisation (RTO) or a school in partnership with a RTO.

Co-ordination at the national level

At a national level there are various consultative arrangements that exist, such as the Council of Australian Governments (COAG) and the Ministerial Council on Education, Early Childhood Development and Youth Affairs (MCEECDYA), to ensure that governments across the country can work together on shared priorities and agree to national initiatives. COAG and MCEECDYA in particular are the main pillars of the current national agenda on education.

Council of Australian Governments (COAG)

COAG, established in 1992, is the peak intergovernmental forum in Australia. COAG comprises the Prime Minister, state Premiers, territory Chief Ministers and the President of the Australian Local Government Association (ALGA). The role of COAG is to initiate, develop and monitor the implementation of policy reforms that are of national significance and which require co-operative action by Australian governments. Through COAG, Australian governments have agreed to a shared policy framework to work towards COAG's educational targets and outcomes.

Ministerial Council on Education, Early Childhood Development and Youth Affairs (MCEECDYA)

MCEECDYA is the principal forum for developing national priorities and strategies for schooling. Membership of the Council comprises state, territory, Australian Government and New Zealand Ministers with responsibility for the portfolios of school education, early childhood development and youth affairs. Functions of the Council include co-ordination of strategic policy at the national level, negotiation and development of national agreements on shared objectives and interests (including principles for Australian Government/state relations) in the Council's areas of responsibility.

Australian Curriculum, Assessment and Reporting Authority (ACARA)

ACARA, established by MCEECDYA, takes responsibility for managing the creation and implementation of a national curriculum, national student assessment and reporting nationally on school education outcomes (including school and system performance). ACARA is jointly funded by the states and territories and the Australian Government and commenced operation in May 2009. The establishment of ACARA brings together the management of curriculum, assessment and reporting for the first time in Australia at the national level and aims to provide a central mechanism through which all Australian Governments can drive national education priorities.

Australian Institute of Teaching and School Leadership (AITSL)

AITSL, established in January 2010, provides national leadership for the Australian Government and state and territory governments in promoting excellence in the profession of teaching and school leadership. AITSL's role includes: developing and maintaining rigorous national standards for teaching and school leadership; implementing an agreed system of national accreditation of teachers based on these standards; administering annual prestigious national awards for teachers and school leaders; undertaking and engaging with international research and innovative developments in best practice; fostering and driving high-quality professional development for teachers and school leaders; working collaboratively across jurisdictions and sectors; and engaging with key professional bodies and stakeholders.

Australasian Curriculum, Assessment and Certification Authorities (ACACA)

ACACA is the national body for the chief executives of the statutory bodies in the Australian states and territories and in New Zealand responsible for certificates of senior secondary education. ACACA provides a national means for monitoring and enhancing developments in senior secondary curriculum and certification. ACACA provides advice on curriculum, assessment and certification matters, including matters of national concern for senior secondary education.

Funding

Public funding for schooling in Australian is shared between levels of government. Private contributions, mainly in the form of fees from parents, support the operation of schools, particularly for non-government schools. In general, state and territory governments provide the majority of recurrent funding to government schools, and the Australian Government is the primary source of public funding for the non-government schooling sector.

In 2007-08, the Australian Government and the state and territory governments' contributions to school education amounted to AUD 36.4 billion, of which AUD 28.8 billion (79%) was expended on government schools and AUD 7.6 billion (21%) expended in non-government schools. For government schools, state and territory governments provided 91.4% of total government recurrent expenditure in 2007-08 and the Australian Government provided 8.6%. For non-government schools, the Australian Government contributed 72.1% of public recurrent expenditure and state and territories 27.9%.

Student learning outcomes considerably above the OECD average but showing some decline

Student learning outcomes in Australia are very good by international standards even if there is evidence of some decline in the last decade. In 2009, achievement levels of Australian students in the OECD's Programme for International Student Assessment (PISA) were significantly above the OECD average in each of the assessment domains – reading, mathematics and science (OECD, 2010a). However, trend analyses of PISA results have raised concerns about a decline in student learning outcomes.

In PISA 2009, the main focus was on reading literacy. The performance of Australian 15-year-olds in reading was significantly above the OECD average – only six countries scored significantly higher than Australia. However, results significantly decreased since

the first PISA study in 2000 (OECD, 2010b) – Australia is among the five OECD countries for which performance declined significantly between 2000 and 2009. The mean score for Australian students in PISA 2000 was 528 points, compared to 515 for PISA 2009. In terms of the proficiency levels, the proportion of students who achieved Level 5 or 6 declined significantly from 18% in PISA 2000 to 13% in PISA 2009. At the lower end of the reading literacy proficiency scale, 12% of students failed to reach Level 2 in PISA 2000 compared to 14% in PISA 2009 (Thomson *et al.*, 2011).

The results of Australian 15-year-olds in mathematics are also considerably above the OECD average – only 12 countries significantly outperformed Australia. However, the PISA 2009 results indicated a fall in test scores in comparison to the PISA in-depth assessment of mathematics in 2003 (OECD, 2010b). In PISA 2009, the average mathematics score was 514 points, ten points lower than it was in 2003 – representing a statistically significant decline in mathematical literacy (Thomson *et al.*, 2011). Science results of Australian 15-year-olds were also above the OECD average in 2009 – only six countries scored significantly higher than Australia and in this assessment area there was no significant change in the average scores between 2006 and 2009 (Thomson *et al.*, 2011). Results from the IEA's Trends in International Mathematics and Science Study (TIMSS) further showed a significant decline in mathematics for eighth-grade students between 1995 and 2007, although there was no significant change between 2003 and 2007 (Thomson *et al.*, 2009). This was in contrast to the significant increase in mathematics performance for fourth-grade students between TIMSS 1995 and TIMSS 2007. In the TIMSS science assessment, there was no significant change in performance between 1995 and 2007 for fourth graders, while for eighth graders Australia had an increase between 1995 and 2003 that was balanced out by a decrease in 2007 (Thomson *et al.*, 2009).

The variation in performance between high- and low-performing students in Australia was higher than the OECD average in reading and science and similar to that found for the OECD as a whole in mathematics in PISA 2009 (Thomson *et al.*, 2011). In reading literacy, the gap between students in the highest and lowest socio-economic quartile is equivalent to more than one proficiency level or almost three full years of schooling (Thomson *et al.*, 2011). The performance of Indigenous students is considerably below the Australian average. For instance, Indigenous students scored 82 points lower, on average, than non-Indigenous students in reading literacy – this difference equates to more than one proficiency level or more than two full years of schooling (Thomson *et al.*, 2011). No statistically significant difference was observed in variation in student performance in reading between 2000 and 2009 (OECD, 2010c). Variations in student reading performance can mostly be found within schools (OECD, 2010c). Such variation significantly decreased between 2000 and 2009 but remains above the OECD average. The between-school variation of reading performance in Australia remains lower than the OECD average, which seems to indicate that the specific school a student attends has only a modest impact on how the student performs (OECD, 2010c).

Regarding the PISA relationship between socio-economic background and performance (*i.e.* between the PISA index of economic, social and cultural status and the performance of 15-year-olds), the following indications emerge: *(i)* Australia is not statistically different from the OECD average in terms of the percentage of variance in student performance explained by student socio-economic background (strength of the socio-economic gradient), *i.e.* the likelihood of disadvantaged students performing at levels similar to those of their advantaged peers is around the OECD average; and *(ii)* Australia is significantly above the OECD average in terms of the score point

difference associated with one unit increase in the PISA index of economic, social and cultural status (slope of the socio-economic gradient) (OECD, 2010c) – and there was no significant change between 2000 and 2009 in this indicator.

Main policy developments: the national agenda for education

A common framework for reform: the National Education Agreement

In 2008 COAG set out its national reform agenda – the COAG National Productivity Agenda – with the goals to boost productivity, workforce participation and geographic mobility, and support wider objectives of better services for the community, social inclusion, closing the gap on Indigenous disadvantage and environmental sustainability. In the area of education, this resulted in all governments agreeing to a common framework for reform in education through the National Education Agreement (NEA) (COAG, 2008). The NEA articulates the roles and responsibilities of the Australian Government and the states and territories. It does not impose input controls on how state and territory governments spend Australian Government funding, as has historically been the case, but instead it focuses on outcomes. Under the NEA, state and territory governments are responsible for developing policy, delivering services, monitoring and reviewing performance of individual schools and regulating schools so as to work towards national objectives and achievement of outcomes compatible with local circumstances and priorities.

The three major reform priorities set by the Australian Government are raising the quality of teaching in schools, ensuring all students are benefiting from schooling, especially in disadvantaged communities, and improving transparency and accountability of schools and school systems at all levels. The new framework – COAG National Productivity Agenda – Schools – includes a set of aspirations, outcomes, progress measures and future policy directions to guide education reform across the country, including a strong focus on Indigenous and also low socio-economic status students in order to lift outcomes for these groups (see Table 2.2). These are articulated through the NEA.

The NEA provides the vehicle through which the Australian Government provides funding to states and territories for government schools. Through this agreement, all Australian schools have an obligation to meet a common set of national school performance and reporting requirements (funding for non-government schools is appropriated separately through the Schools Assistance Act 2008, which entails similar requirements). The new framework for financial relations establishes clear roles and responsibilities for each level of government, sets outcome measures on the states and territories accompanied by transparent accountability reducing Australian Government prescriptions on service delivery by states and territories. As part of inter-government financial arrangements, the Australian Government also provides additional funding through National Partnership arrangements which provide financial support to achieve specifically agreed outcomes, such as computers in schools, or in the form of reward payments for implementation of reforms (see below). The Australian Government is currently conducting a review of funding for schooling which will conclude in 2011. The review is expected to be extensive and will inform Government decisions beyond 2013, the year the current non-government school funding arrangements conclude.

Table 1.2 COAG's National Productivity Agenda – Schools

Aspirations	That all Australian school students acquire the knowledge and skills to participate effectively in society and employment in a globalised economy					
Outcomes	All children are engaged in and benefiting from schooling	Young people are meeting basic literacy and numeracy standards, and overall levels of literacy and numeracy achievement are improving	Australian students excel by international standards	Schooling promotes the social inclusion and reduces the educational disadvantage of children, especially Indigenous children	Young people make a successful transition from school to work and further study	
Indicative progress measures	o Proportion of children enrolled in and attending school o Literacy and numeracy achievement of Year 3, 5, 7 and 9 students in national testing o Proportion of students in top and bottom levels of performance in international testing (*e.g.* PISA, TIMSS) o Proportion of the 19-year-old population having attained at least a Year 12 or equivalent or AQF Certificate II o Proportion of young people participating in post-school education or training six months after school o Proportion of 18-24 year-olds engaged in full-time employment, education or training at or above Certificate level III					
COAG targets	o Lift the Year 12 or equivalent attainment rate to 90% by 2020 o Halve the gap for Indigenous students in reading, writing and numeracy within a decade o At least halve the gap for Indigenous students in Year 12 or equivalent attainment rates by 2020					
The Australian Government election commitments 2007	o All Year 9-12 students have access to a computer and teachers are trained appropriately o All secondary schools have access to Trades Training Centres delivering high-quality industry-recognised training at Certificate III level o Australian Curriculum supports world-class teaching in all Australian schools from Foundation to Year 12, including literacy and numeracy standards o Asian Languages: Increase number of qualified language teachers and develop national curriculum for advanced students					
Policy directions	Improving teacher and school leader quality	High standards and expectations	Greater accountability and better directed resources	Modern, world-class teaching and learning environments including ICT	Integrated strategies for low SES school communities	Boosting parental engagement

Source: Reproduced from Australian Government (2010).

Melbourne Declaration on Educational Goals for Young Australians

The Melbourne Declaration on Educational Goals for Young Australians, released in December 2008, and agreed to by all education ministers through MCEECDYA, articulates future directions and aspirations for Australian schooling (MCEETYA, 2008). It sets young Australians at the centre of the agenda for educational goals and provides a framework for developing curriculum and assessment. The Melbourne Declaration has two overarching goals for schooling in Australia:

- Australian schooling promotes equity and excellence; and
- All young Australians become successful learners, confident and creative individuals, and active and informed citizens.

The national goals for schooling are supported by the MCEECDYA Four-Year Plan 2009–2012 (MCEETYA, 2009), which was endorsed by all Australian education ministers in March 2009. The plan is closely aligned with the COAG agreements. It outlines the key strategies and initiatives Australian governments will undertake in the

following eight inter-related areas in order to support the achievement of the educational goals outlined in the Melbourne Declaration:

- Developing stronger partnerships;
- Supporting quality teaching and school leadership;
- Strengthening early childhood education;
- Enhancing middle years development;
- Supporting senior years of schooling and youth transitions;
- Promoting world-class curriculum and assessment;
- Improving educational outcomes for Indigenous youth and disadvantaged young Australians, especially those from low socio-economic backgrounds; and
- Strengthening accountability and transparency.

Australian Curriculum

Up until recently, states and territories have been responsible for setting the curriculum and achievement standards for their state or territory, through Boards of Studies or relevant authorities. Curriculum content and achievement standards are mandated under state and territory regulations, usually in the form of Education Acts which apply to all schools registered within each state or territory. This ensures both government and non-government schools are required to implement and follow the curriculum and standards. In 2008, all Australian education ministers committed to the development and implementation of a national curriculum for Foundation (year of schooling before Year 1) to Year 12, beginning with the learning areas of English, mathematics, science and history. The Australian Curriculum in the four initial learning areas from Foundation to Year 10 was endorsed by all education ministers in December 2010 and will begin to be implemented from 2011 with substantial implementation in all states and territories by 2013. The Australian Curriculum can be viewed at *www.australiancurriculum.edu.au*.

The Australian Curriculum in the initial learning areas for the senior secondary years will follow in 2012. The next phase of work will involve the development of an Australian Curriculum in languages, geography and the arts, while future phases will focus on health and physical education, information and communication technology, design and technology, economics, business, and civics and citizenship. ACARA's work in developing the Australian Curriculum is being guided by the Melbourne Declaration on Educational Goals for Young Australians.

The Australian Curriculum provides two key elements: *(i)* Agreement on the curriculum content that all Australian students should be taught (outline of knowledge, skills and understandings for each learning area at each year level); and *(ii)* Explicit advice on the achievement standards that all Australian students should be meeting (depth of understanding, extent of knowledge and sophistication of skill expected of students at each year level). For each learning area, the achievement standards will comprise: *(i)* A description of the quality of learning expected; and *(ii)* A set of work samples that illustrate the described quality of learning.

Aboriginal and Torres Strait Islander Education Action Plan 2010-2014

The Aboriginal and Torres Strait Islander Education Action Plan (the Action Plan) outlines how governments will work together to achieve nationally agreed targets to close the gaps between the educational outcomes of Aboriginal and Torres Strait Islander peoples and other Australians. The Action Plan supports the goals of the National Aboriginal and Torres Strait Islander Education Policy and the Melbourne Declaration on the Educational Goals of Young Australians. All Australian governments and non-government education authorities have agreed to progress actions outlined in the Action Plan. The Action Plan aims to ensure that across-government commitments to introduce substantial structural and innovative reforms in early childhood education, schooling and youth accelerate improvements to the outcomes of Aboriginal and Torres Strait Islander students. The Action Plan also complements these mainstream initiatives as outlined in the National Education Agreement and national partnership agreements with a number of new targeted activities.

The Action Plan identifies national, systemic and local actions in six strategic priority domains that evidence shows will contribute to improved outcomes in Aboriginal and Torres Strait Islander education including:

- Readiness for school;
- Engagement and connections;
- Attendance;
- Literacy and numeracy;
- Leadership, quality teaching and workforce development; and
- Pathways to real post-school options.

The accountability framework established through the Action Plan clearly attributes responsibility for targets, performance indicators and actions. The Plan is supported by annual public reporting and an ongoing evaluation strategy to ensure actions beyond 2014 are informed by the best available evidence.

National Partnerships

In addition to funding from the NEA, the Australian Government is providing significant additional funding through collaborative new National Partnerships with states and territories. The following National Partnerships were announced in November 2008 or subsequently by COAG or have been deemed to be part of the National Partnerships (NPs):

- Smarter Schools National Partnership for Literacy and Numeracy;
- Smarter Schools National Partnership for Low Socio-Economic Status (SES) School Communities;
- Smarter Schools National Partnership for Improving Teacher Quality;
- National Partnership on Youth Attainment and Transitions;
- Digital Education National Partnership;
- Trade Training Centres National Partnership; and

- Building the Education Revolution (part of the National Building – Economic Stimulus Plan).

Bilateral agreements between the Australian Government and each jurisdiction have been developed and the detail of Implementation Plans under each of the "Smarter Schools" NPs and the NP on Youth Attainment and Transitions have been negotiated. The non-government school sector is participating in relevant NPs. States and territories have worked with non-government schools and system authorities to determine funding arrangements through bilateral agreements.

The Australian Government is investing AUD 2.59 billion through the three Smarter Schools National Partnerships:

- Literacy and Numeracy NP: supports states and territories to implement evidence-based practices that will deliver sustained improvement in literacy and numeracy outcomes for all students, especially those who are most in need of support.
- Improving Teacher Quality NP: is aimed at developing effective workforce planning and supporting structures to identify teaching performance and to reward quality teaching at the national level.
- Low SES School Communities NP: aims to better support student learning needs and well-being by facilitating education reform activities in up to 1 700 low socio-economic status schools across the country.

States and territories have agreed to share and collaborate on key reforms under the three Smarter Schools National Partnerships. There are six national projects receiving Australian Government funding to support this national collaboration:

- School performance improvement frameworks;
- Innovative strategies for small and remote schools;
- Parental engagement in schooling in low SES communities;
- Extended service models in schools;
- Literacy and numeracy diagnostic tools;
- School leadership development strategies.

National Assessment Program (NAP)

The National Assessment Program (NAP) is an ongoing programme of assessments, agreed by MCEECDYA, to monitor progress towards the Educational Goals for Young Australians and to support ongoing evaluation of the national education system. The NAP encompasses the National Assessment Program – Literacy and Numeracy (NAPLAN) and three-yearly sample assessments in science literacy, civics and citizenship, and information and communication technology (ICT) literacy. Australia's participation in international assessments – Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS) – are also part of the NAP.

References

Australian Government (2010), *Country Background Report for Australia*, prepared for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, Department of Education, Employment and Workplace Relations, Canberra, available from *www.oecd.org/edu/evaluationpolicy*.

COAG (Council of Australian Governments) (2008), *National Education Agreement: Intergovernmental Agreement on Federal Financial Relations*, Sydney, *www.coag.gov.au/intergov_agreements/federal_financial_relations/index.cfm*.

MCEETYA (2008), *Melbourne Declaration on Educational Goals for Young Australians*, Ministerial Council on Education, Employment, Training and Youth Affairs, Canberra.

MCEETYA (2009), *MCEETYA Four-Year Plan 2009-2012 – A Companion Document for the Melbourne Declaration on Education Goals for Young Australians*, Ministerial Council on Education, Employment, Training and Youth Affairs.

OECD (2010a), *PISA 2009 Results: What Students Know and Can Do, Student Performance in Reading, Mathematics and Science*, Volume I, OECD, Paris.

OECD (2010b), *PISA 2009 Results: Learning Trends: Changes in Student Performance since 2000*, Volume V, OECD, Paris.

OECD (2010c), *PISA 2009 Results: Overcoming Social Background: Equity in Learning Opportunities and Outcomes*, Volume II, OECD, Paris.

Thomson, S., N. Wernert, C. Underwood and M. Nicholas (2009), *TIMSS 2007: Taking a Closer Look at Mathematics and Science in Australia*, Australian Council for Education Research, Camberwell, Victoria, *www.acer.edu.au/documents/TIMSS_2007-AustraliaFullReport.pdf*.

Thomson, S., L. De Bortoli, M. Nicholas, K. Hillman and S. Buckley (2011), *Challenges for Australian Education: Results from PISA 2009: The PISA 2009 Assessment of Students' Reading, Mathematical and Scientific Literacy*, Australian Council for Education Research, Camberwell, Victoria, *www.acer.edu.au/documents/PISA-2009-Report.pdf*.

Chapter 2

The evaluation and assessment framework

Evaluation and assessment in Australia operates at four key levels: (i) National and systemic (state, territory or non-government system) – namely through the National Assessment Program and state- and territory-based assessments; (ii) School – a variety of forms of school evaluation typically in the context of a School Performance Improvement Framework; (iii) Teacher – through registration processes, performance management, and Advanced Skills Teaching positions; and (iv) Student – with instruments ranging from national standardised tests to ongoing daily formative assessment in the classroom. The overall evaluation and assessment framework appears as highly sophisticated and well conceptualised, especially at its top level (national and systemic levels). Particularly positive characteristics of the framework include the national educational goals as a solid reference point; the strong capability at the national level to steer evaluation and assessment; a focus on student outcomes; a coherent system of assessments for learning; a good structure to integrate accountability and improvement; and the commitment to transparency. Priorities for future policy development include establishing national strategies for strengthening the linkages to classroom practice; promoting greater national consistency while giving room for local diversity; improving the integration of the non-governmental sector in the overall framework; further developing some articulations within the overall framework and sustaining efforts to improve capacity for evaluation and assessment.

This chapter looks at the overall framework[1] for evaluation and assessment systems in Australia, *i.e.* its various components such as student assessment, teacher appraisal, school evaluation and system evaluation, the coherence of the whole as well as the articulation between the different components. Following this overview, the succeeding chapters (3-6) will analyse the issues relevant to each individual component in more depth.

This report differentiates between the terms "assessment", "appraisal" and "evaluation". The term "assessment" is used to refer to judgments on individual student progress and achievement of learning goals. It covers classroom-based assessments as well as large-scale, external tests and examinations. The term "appraisal" is used to refer to judgements on the performance of school-level professionals, *e.g.* teachers. Finally, the term "evaluation" is used to refer to judgments on the effectiveness of schools, school systems and policies. The term "review" is also used in the context of school evaluation.

Context and features

Objectives

Evaluation and assessment in Australia operates at four key levels: national and systemic (state, territory or non-government system); school; teacher and student. According to the Australian Country Background Report (Australian Government, 2010), components of the overall evaluation and assessment framework "at both the national and state and territory levels are generally outcomes focused, put students at the centre and are strongly influenced by the goal to ensure all Australian school students receive "quality education" – that they acquire the knowledge and skills to participate effectively in society and employment in a globalised economy."

Evaluation and assessment mechanisms provide a basis for measuring and reporting against the relative performance of schools, systems and students, and for assessing how effectively education is being delivered to students in Australia. Evaluation and assessment also identifies strengths and weaknesses of systems, schools, teachers and students which inform areas for improvement. The overall framework exists in an environment where there is a growing trend of reporting and accountability for all governments and educational institutions. Evaluation and assessment in Australia informs budgetary discussions, resource allocation decisions, curriculum, planning, reporting and performance management.

Components of the overall framework

Main components

In a nutshell, the Australian overall framework for evaluation and assessment can be described as consisting of the following four main components (see Table 2.1):

- **Student assessment**. Student performance in Australia is assessed by a wide range of instruments, ranging from national standardised tests to ongoing daily formative assessment in the classroom. At the national level, both full-cohort and national sample assessments of Australian students are conducted, the results from which are used as key performance measures towards national goals. At the

system level, many state and territory governments administer testing with both diagnostic and monitoring purposes. States and territories are also responsible for externally-based summative assessment, in particular in view of assessing students for secondary education certification. At the school level, student assessment plays the key role in informing schools and teachers about students' individual achievement through teacher-based summative and formative assessments.

- **Teacher appraisal**. Procedures vary across states, territories and school sectors but, in addition to probationary processes, typically occur in three specific instances: (1) to gain and maintain registration/accreditation to teach within the state or territory (with procedures mostly school-based which currently evaluate against jurisdiction-specific teaching standards); (2) as part of the employer's performance management processes (in general an annual process internal to the school and typically linked to school improvement frameworks); and (3) to gain promotion positions in schools in recognition of quality teaching performance (Advanced Skills Teaching positions).
- **School evaluation**. Australia has a variety of forms of school evaluation in place, each of which derives from the particular circumstances and traditions of the state, territory and school sector within which it has developed. There are two main forms of evaluation: school self-evaluation and school external performance review. This is represented as a sequence of activities which begins with self-reflection by the school and proceeds through a planning, reporting and review process which both satisfies external requirements and is an engine of school improvement. The precise nature of school self-evaluation varies across jurisdictions but it is generally seen as contributing directly to developing or monitoring school plans. External school reviews vary widely across jurisdictions and in government schools work within a clear state or territory policy – typically a School Performance Improvement Framework – and are organised and staffed by relevant state government departments.
- **System evaluation**. Monitoring progress towards educational goals is a priority both at the national and systemic levels. This is accomplished namely through the National Assessment Program and state- and territory-based assessments. The monitoring system also includes a *Measurement Framework for National Key Performance Measures* as well as data and surveys at the systemic level. The strategy draws considerably on public reporting of the progress and performance of Australian students and schools through instruments such as the *My School* website, the *National Report on Schooling in Australia*, COAG Reform Council Reports, Report on Government Services in addition to system-level analyses organised through independent reviews.

Common elements

Although evaluation and assessment in Australia are specifically designed for the system or part of the system in which they operate, evaluation and assessment at all four levels include the following common elements, as described in Table 2.1:

- **Strategic goals**. Setting strategic educational goals provides a basis for policy development and curriculum as well as measurement and reporting. It is these goals that sit at the centre and form the basis for evaluation and assessment;

school and system outcomes are evaluated and assessed against these goals to determine relative performance and further plan and identify policy priorities.

- **Reference standards**. These refer to the standards against which a specific evaluation or assessment is undertaken and are typically aligned with educational goals set at the system level. They vary according to the level at which the evaluation or assessment is conducted.
- **Evaluation/Assessment**. These refer to the types of evaluation and assessment – range of instruments and information sources – used at the different levels in the overall framework. These seek to measure progress towards achievement of goals and standards at the system, school, teacher and student levels.
- **Reporting**. These refer to the reporting practices associated with evaluation and assessment at the different levels of the overall framework. They provide an analytical tool to inform further goals setting and/or the planning and identification of further policy priorities as well as ensuring accountability.

Table 2.1 specifies the strategic goals, reference standards, evaluation/assessment, and reporting practices at each level of the overall evaluation and assessment framework.

An important feature of evaluation and assessment in Australia is that the National Education Agreement (NEA) includes a set of reporting requirements for all Australian schools, including those which are part of the non-government sector. The five basic requirements are:

- **National testing**. All schools are required to participate in the assessments which are part of the National Assessment Program.
- **National reporting**. All schools and system authorities must participate in preparing national reports on the outcomes of schooling.
- **(National) individual school information**. All schools are required to provide individual school information on the school's context, capacity (including school income) and outcomes, to enable nationally comparable information about each school to be made publicly available (this information is published on the *My School* website).
- **Reporting to parents**. Requires that student reports to parents use plain language, give an accurate assessment of progress, and include assessment of achievement against national standards and relative to the student's peer group.
- **Publication of information relating to schools (school annual reports)**. Aimed at parents and the community, schools must publish an annual report online which includes contextual information about the school; key outcomes; information on satisfaction; and income by funding source.

Table 2.1 Levels of evaluation and assessment in Australia

Level		Strategic goals	Reference standards	Evaluation/Assessment	Reporting
System	**National**	• COAG National Productivity Agenda • Melbourne Declaration on Educational Goals for Young Australians Operationalised through: • National Education Agreement • National Partnerships • Aboriginal and Torres Strait Islander Education Action Plan • Australian Curriculum	• Student learning objectives • Priority areas / educational targets	• National Assessment Program (NAPLAN, sample assessments, international assessments) • *Measurement Framework for National Key Performance Measures* • Independent reviews	• COAG Reform Council Reports • *National Report on Schooling in Australia* • *My School* website • Report on Government Services • Reports from Independent Reviews
	Systemic (state and territory, Catholic, Independent)	• State, territory or other systemic strategic plans Operationalised through: • Departmental / system plans • Programmes / funding initiatives • State / territory curriculum	• Student learning objectives at the national and systemic levels • Priority areas / educational targets at the national and systemic levels	• State- / territory-based assessment programmes • Surveys and other data	• System annual report • School data reporting
School		• Educational goals at the national and systemic levels • School performance improvement framework • School strategic plans	• Student learning objectives at the national and systemic levels • Priority areas / educational targets at the national, systemic and school levels • Action / operational / improvement plans at the school level	• School self-evaluation • School performance review	• School annual report • Performance review reports • School data reporting • *My School* website • Annual improvement / action plans
Teacher		• Educational goals at the national, systemic and school levels • School performance improvement framework • School strategic plans	• Teaching standards at systemic levels (national standards currently being developed)	• Registration / accreditation processes • Performance management processes • Advanced Skills Teaching positions • Probationary period	• Registration status • Pay increment • Annual improvement / action plan • Probation decision • Promotion decision
Student		Educational goals at the national and systemic levels	The Australian Curriculum and state / territory Curriculum	• Teacher-based summative assessment • Classroom-based formative assessment • Externally-based summative assessment (*e.g.* for the senior secondary certificate) • Assessment for certification (secondary education) • Standardised diagnostic assessment • Standardised tests to monitor national objectives (NAPLAN and sample assessments in science, ICT, and civics and citizenship) • Standardised assessment at systemic level • International student assessments	• A-E reporting • Senior Certificate • National, systemic and school-level reporting

Source: Australian Government (2010).

Strengths

Clear national goals for education in Australia which gather wide support

National goals for education which respond to broader social and economic needs

The OECD Review Team formed the impression that there is wide support across the system for the national agenda for education. The objectives and priorities as articulated in both the COAG National Productivity Agenda and the Melbourne Declaration on Educational Goals for Young Australians are well understood by the different stakeholders. The vision is clear and shared across the system. This is a major strength to achieve the alignment of processes and school agents' contributions within the overall evaluation and assessment framework. In addition, educational goals are founded on a rationale which gathers consensus among stakeholders, namely their alignment with broader social and economic goals. Indeed, national educational goals are set in light of their contribution to social cohesion and economic growth, as articulated in COAG's National Productivity Agenda (see Chapter 1). They respond to ample social and economic needs and hence reflect perspectives and views from outside the education sector. Another positive feature is that educational goals are established in a way to ensure their continuity in the longer term. For instance they are informed by the research and analysis undertaken by the Productivity Commission[2] which plays a major role in guiding policy across a range of economic and social issues in Australia within a long-term perspective, including with the monitoring of education outcomes (*e.g.* through the production of the *Report on Government Services*) (see Chapter 6). The clear and widely supported national education goals provide a solid reference point on which to build evaluation and assessment practices.

States and territories increasingly align their goals with the national agenda

The set of national goals for education in Australia were collaboratively agreed on by states and territories. Ministers in 2009 set out a four-year plan to work towards these national goals, detailing agreed strategies in the eight commitments to action included in the Melbourne Declaration (MCEETYA, 2008) (see Chapter 1). Accordingly, strategic plans at the state and territory level are increasingly aligned with the national agenda. For example, the Queensland Strategic Plan 2009-13 includes the goals "laying strong educational foundations" and "developing skills for the economy", two of the four priorities in Western Australia include "attendance" and "literacy and numeracy", and in South Australia the three goals for children and students are: "strong beginnings for all children; excellence in learning; engagement and well-being". In Victoria, the outcomes for government schools are in alignment with the COAG key outcomes: "All children are engaged in and benefiting from schooling; children are meeting expected literacy and numeracy standards, and overall levels of literacy and numeracy are improving; Victorian students excel by national and international standards" (Department of Education and Early Childhood Development, 2010). In some cases the goals are explicitly linked to the overall national agenda. For example, in the Australian Capital Territory one of the performance indicators is "implementing COAG reforms in education, skills and early childhood development" and New South Wales aims "to exercise strong leadership in Australian education and training through innovation and by shaping national policy and reform". Therefore, the national goals

have influenced system-level goals, but may be given different emphasis depending on the major challenges within each system.

As part of the NEA, National Partnerships have been developed to promote the education reform agenda within specific target areas and suitably serve as the main funding instrument to support implementation. These National Partnerships provide an incentive mechanism for encouraging greater consistency in approach by funding joint working between states and territories to promote innovation and the spread of best practice.

Equity is of central importance

Equity is at the core of the national goals for education, and national reporting on education pays careful attention to different measures of equity, including gender, Indigenous groups, geographic location and students with a language background other than English (LBOTE). Equity has been given more prominence in general government reporting since 2004 when it was put on the same level as "efficiency" and "effectiveness" in the Report on Government Services' general performance indicator framework, with indicators on equity of access (output) and equity of outcomes (see Box 2.1).

Box 2.1 Equity within the performance framework

High priority is accorded to equity in the current national agenda for education. One of the two overarching goals in the Melbourne Declaration is that "Australian schooling promotes equity and excellence". Two of the three COAG targets are to reduce educational disadvantage for Indigenous students, namely to halve the gap in their reading, writing and numeracy achievement levels and in their educational attainment at Year 12 or equivalent by 2020.

Accordingly, concerns for equity in education have been given increased importance in national reporting:

- The annual *Report on Government Services* produced by the Productivity Commission places equity alongside effectiveness and efficiency at the heart of the performance measurement framework. Equity is defined along the tradition of economic literature as both:
  - Horizontal – "when services are equally accessible to everyone in the community with a similar level of need".
  - Vertical – "when services account for the special needs of particular groups in the community and adjust aspects of service delivery to suit these needs". "Special needs" in this light could include geographical, cultural or other factors that impede access to a standard government service.
- The prominence of education indicators among the headline indicators in the *Overcoming Indigenous Disadvantage: Key Indicators* report series has also increased from a single indicator in the 2007 report on retention and attainment in Years 10 and 12, to three COAG targets (early childhood education; reading, writing and numeracy; Year 12 attainment) in the 2009 report (although there are limited data on Indigenous preschool participation).
- ACARA reporting of NAPLAN results includes the proportion of students who were absent, exempt or withdrawn from the tests. Such transparency aims to promote equity of participation in NAPLAN by all eligible students. NAPLAN results are also reported by: sex; Indigenous status; language background other than English (LBOTE); geographic location; Indigenous geographic location; non-Indigenous geographic location; parental education; and parental occupation.
- National analysis of Australian results from the PISA surveys has paid considerable attention to equity issues. Results for Indigenous students are reported in the main national reports of initial PISA results. For example, there have been a series of analytical reports on results for Indigenous students, including a report analysing the influence of different contextual factors over performance of Indigenous students in PISA 2000, 2003 and 2006 (De Bortoli and Thomson, 2010). Further, using the PISA 2000 results, researchers identified a strong association between the geographic location of students' schools and student outcomes and different school factors (Cresswell and Underwood, 2004).

Source: Steering Committee for the Review of Government Service Provision (2009).

Sophisticated and well-conceptualised overall evaluation and assessment framework which builds on the national agenda for education

The national agenda for education has granted the opportunity to conceptualise an overall evaluation and assessment framework at the national level. This has been achieved through the development of goals, monitoring and reporting at the national level as well as mechanisms to articulate national objectives with jurisdiction-level goals and priorities. To the Review Team, the overall evaluation and assessment framework appears as highly sophisticated and well conceptualised, especially at its top level (national and

systemic levels). Particularly positive characteristics of the evaluation and assessment framework include:

- **The national educational goals are a solid reference point** (see above) and instruments such as the National Education Agreement, the Australian Curriculum, the Aboriginal and Torres Strait Islander Education Action Plan and the National Partnerships warrant their application across the system. These have the potential to promote greater coherence in the interpretation of goals for student learning and greater consistency of evaluation and assessment practices within and across schools while leaving enough room for adaptation to local needs.
- **Strong capability at the national level to steer evaluation and assessment**: In addition to the direction and guidance provided by COAG and MCEECDYA, the establishment of ACARA and AITSL provides important leadership in monitoring and reporting on student outcomes, and in ensuring good standards of teaching and school leadership. Both bodies build on high-level expertise and foster the development of skills for evaluation and assessment across the system.
- **A focus on student outcomes**: Evaluation and assessment in Australia focuses on improving student outcomes and achieving student learning objectives. This is reflected in the priorities for national monitoring (in particular the National Assessment Program), the significance of evidence on student performance for school and teacher evaluation, and the importance of reporting publicly on student results.
- **A coherent system of assessments for learning**: The range of provisions for the assessment of student learning (NAPLAN, sample-based assessments, international assessments, A-E ratings, senior secondary and VET certificates) has the potential to provide a comprehensive picture of student performance relative to Australia's goals for student learning (see Chapter 3).
- **A structure to integrate accountability and improvement**: The overall evaluation and assessment framework includes elements to accomplish both the accountability and improvement functions at all levels of the system (*e.g.* formative *vs.* summative assessment for students; professional development for teachers *vs.* promotion decisions following teacher appraisal; data reporting *vs.* improvement action plans for schools) and provides a structure which can potentially integrate these two functions.
- **The commitment to transparency**: The overall evaluation and assessment framework is strengthened by a high level of transparency in monitoring and publishing results. Reporting, as one of the main functions of the evaluation and assessment framework, receives high priority as reflected in the requirements at several levels: system level (*e.g.* COAG Reform Council Reports, Report on Government Services); school level (*My School* website, School Annual Report, Performance Review reports); and student level (A-E Reporting).

The principle of evidence-based policy is well established

The principle of informing policies and evaluation and assessment practices with evidence from research is well established in Australia. The concern of evaluating policies and identifying best practice exists across the system, including at the level of practitioners in Australian schools. For instance, within the Department of Education,

Employment and Workplace Relations (DEEWR), the Strategic Information Management and Research (SIMR) Committee works to ensure a strategic approach to DEEWR's research, analysis and evaluation, and information management activities in view of supporting the provision of evidence-based policy advice to the Minister for Education and other parliamentary officers. In addition, DEEWR commissions a variety of research studies and promotes evaluation activities and data collections to inform its policy advice. Institutions such as the Australian Council for Educational Research (ACER) and the National Centre for Vocational Education Research (NCVER) are among the main contributors of educational research which informs policy development. Evidence-based policy is also at the heart of the work of the Productivity Commission, COAG and MCEECDYA, as illustrated by the focus of a recent Roundtable promoted by the Productivity Commission: Strengthening Evidence-based Policy in the Australian Federation (Productivity Commission, 2010).

The same approach is followed at the jurisdiction level. Education Departments within jurisdictions have dedicated units to promote research in education and analyse the implications of research for policy development. This is also common in state and territory statutory bodies such as curriculum and assessment bodies or teaching registration authorities. At the national level, AITSL has as one of its roles "undertaking and engaging with international research and innovative developments in best practice" while ACARA informs its work with evidence on best practices in the areas of curriculum development, assessment and reporting.

There are good bases for sound knowledge management within the overall evaluation and assessment framework

The overall evaluation and assessment framework places great emphasis on the production of data and information on the results it creates and their subsequent use for public information, policy planning and the improvement of practices across the system. This is accompanied by sustained efforts to develop coherent information management systems to make the best use of the evidence generated by evaluation and assessment procedures across the system. This is visible, for instance, in the existence of standard frameworks both for reporting key performance measures (the *Measurement Framework for National Key Performance Measures*) and for general government sector reporting (the *Report on Government Services' Performance Indicator Framework*); the standardised Australian Bureau of Statistics National Schools Statistics Collection (NSSC); and the nationally comparable data on student outcomes (through the National Assessment Program). These entail the establishment of protocols to harmonise, standardise, and share the data among key stakeholders.

Some jurisdictions have also developed sophisticated data information systems – collection of data on students, teachers, schools, and their performance over time. Among the best examples are the School Measurement, Assessment and Reporting Toolkit (SMART) developed by New South Wales and the Ultranet developed by Victoria (see Box 2.2). These notable initiatives have the potential to assist teachers in the instruction of their students, provide quick feedback to school agents, serve as a platform to post relevant instructional material to support teachers and improve knowledge management, operate as a network to connect teachers and schools with similar concerns, and create a better data infrastructure for educational research. In addition, schools' data management systems to track progress of individual students are also common in Australian schools. This means that the development of individual students is tracked over time and that such

information can be shared among teachers or with a student's next school (see also Chapter 6).

Box 2.2 Data information systems in New South Wales and Victoria

New South Wales SMART system

The NSW Department of Education and Training has developed a sophisticated tool for data analysis in the form of the School Measurement, Assessment and Reporting Toolkit (SMART). This provides diagnostic information on NAPLAN, ESSA (a Year 8 NSW based science test) and the NSW School Certificate and Higher School certificate examinations. This information, together with information from school-based assessment activities, provides a wealth of objective diagnostic information to which teachers can respond. The SMART system is an example of how digital technology can assist in effectively using data and is now also used in the Australian Capital Territory and South Australia.

Analysis of educational outcomes and processes in NSW can be undertaken at many levels, from individual students, to groups of students, cohorts, schools and the system as a whole. The SMART package allows educators to identify areas for improvement as well as strengths in student performance. SMART also provides support through specific teaching strategies designed to improve student outcomes. SMART includes a number of functionalities intended to analyse NAPLAN results in-depth (see Table 18 in Australian Government, 2010).

For more information: *www.schools.nsw.edu.au/learning/7-12assessments/smart/index.php*

Source: Australian Government (2010).

Victoria Ultranet system

The Ultranet is a state-wide, secure site that students, parents and teachers can access via the Internet. It provides a new learning space and more opportunities for information sharing across the Victorian government school system. The Ultranet links whole school communities, parents, students and teachers, enabling them to collaborate to improve student learning outcomes in a way not previously possible.

The Ultranet gives parents access to information that will enable them to keep up-to-date with their child's learning progress. This could mean viewing test results, teacher feedback, timetables, homework activities and attendance records. It gives teachers access to learning tools, resources and student information in one place. They can access learning spaces and online tools that extend the classroom; plan and share learning ideas and activities with colleagues across Victoria; access digital resources and collaboratively design and share content with colleagues; and access a rich source of information about each student they teach so they can more easily tailor learning activities to student needs. Students can create an online learning portfolio to keep track of their progress throughout their school life; collaborate with students and teachers using wikis, blogs, polls, message boards and many other web 2.0 tools; access learning tasks, submit work and receive feedback from teachers; and access up-to-date, personalised information about their learning.

For more information: *www.education.vic.gov.au/ultranet*.

Source: Website of Victoria's Department of Education and Early Childhood Development (link provided above).

Challenges

Links to classroom practice are less clearly articulated

The national agenda for education, which is giving rise to greater national consistency, provides a framework of national objectives and establishes clear expectations in relation to the curriculum, teaching standards, student testing, reporting requirements and system monitoring at the top level of the overall evaluation and assessment framework. By contrast, the nature and role of other components of the evaluation and assessment framework such as school evaluation, teacher appraisal or student formative assessment are less well defined in the national reform agenda. The result is a less clear articulation of ways for the national agenda for education to generate improvements in classroom practice through the assessment and evaluation procedures which are closer to the place of learning.

Evaluation and assessment frameworks have no value if they do not lead to the improvement of classroom practice and student learning, and therefore securing effective links to classroom practice is one of the most critical points in designing the evaluation and assessment framework. Examples of potential channels through which the evaluation and assessment framework impacts on classroom practice and which are less well articulated in Australia are: ensuring teaching standards are aligned with student learning objectives; building teacher capacity for student formative assessment; assuring that school-based developmental teacher appraisal is aligned with student learning objectives; and strengthening teachers' ability to assess against A-E standards.

The current focus of a better articulation of evaluation and assessment procedures at the national and systemic levels of the framework also translates into a greater emphasis on the accountability function of evaluation and assessment as the improvement function is more articulated at the local level. The national education agenda has placed considerable investment in establishing national standards, national testing and reporting requirements while it provides considerably less direction and strategy on how to achieve the improvement function of evaluation and assessment. While transparency of information and high-quality data are essential for a well-functioning evaluation and assessment system, there has been comparatively less focus on articulating how the existing data and information should be used for improvement and on ensuring that school agents have the capacity to use the data and feedback made available to them in order to improve their practices. There is no particular national guidance or vision on how the results of evaluation and assessment activities feed back into classroom practice.

Some articulations within the overall evaluation and assessment framework are not sufficiently developed

How the different components have to be interrelated in order to generate complementarities, avoid duplication, and prevent inconsistency of objectives is an important aspect of designing the evaluation and assessment framework. The Review Team noted a number of missing links, or underdeveloped articulations, between different elements of the overall evaluation and assessment framework in Australia. These can be grouped into three distinct sets:

1. Within specific components of the overall evaluation and assessment framework:
   - *Linkages between teacher appraisal and teacher professional development*

     There are some indications that the provision of professional development for teachers is not systematically linked to teacher appraisal (see Chapter 4).
   - *Alignment of teaching standards with student learning objectives*

     This is in the process of being achieved through the development of teaching standards at the national level by AITSL.
   - *Alignment of teaching standards with teaching career structures to reinforce the links between teacher appraisal, professional development and career development*

     This translates into a separation between the definition of skills and competencies at different stages of the career (as reflected in teaching standards) and the roles and responsibilities of teachers in schools (as reflected in career structures) (see Chapter 4).
   - *Articulation between school self-evaluation and external school evaluation*

     There does not seem to be enough reflection about the relative contributions of self-evaluation and external evaluation and the nature of externality for school reviews (see Chapter 5).
   - *Linkages between standardised student testing and student formative assessment*

     There is some lack of clarity about what should be the formative uses of NAPLAN results by teachers (see Chapter 3).
   - *Alignment of A-E ratings with the Australian Curriculum*

     A-E ratings are not currently aligned with the Australian Curriculum and definitions vary across states and territories. There is currently a proposal to establish such link. ACARA is leading this work which is expected to take several years and involve national agreement on definitions (see Chapter 3).
2. Between specific components of the overall evaluation and assessment framework:
   - *Articulation between teacher appraisal, school evaluation and school development*

     This relates to a range of aspects such as: school-based teacher appraisal being validated by school evaluation processes; making the focus of school evaluation on teacher effectiveness systematic across schools; and school development processes exploring links to the evaluation of teaching practice. According to a report by the Grattan Institute, for all areas except for teaching in a multi-cultural setting there was an insignificant correlation between the extent that an aspect of teaching was emphasised in school evaluations and the extent that it was emphasised in the evaluation of teachers in the corresponding school (Grattan Institute, 2010).

– *Articulation between teacher appraisal and student assessment*

There is some lack of clarity about how student results should be taken into account in teacher appraisal and teachers' skills for assessment are not systematically reviewed in teacher appraisal processes (see Chapter 4).

– *Articulation between school evaluation and system evaluation*

Evaluation at the national and systemic levels does seem to make poor use of the information generated by school review processes and there are no mechanisms to ensure the consistency of school evaluation across educational jurisdictions (see Chapter 5).

3. Processes to ensure the consistency of evaluation and assessment procedures:

– *Moderation processes to ensure the appraisal of teachers against teaching standards are consistent across schools and jurisdictions*

This relates to the fact that the extent of externality in teacher appraisal is limited. Teachers are appraised according to local interpretations of common standards with risks of lack of coherence of judgments (see Chapter 4).

– *Moderation processes and capacity development to ensure consistency of teachers' A-E ratings*

Strategies to develop teachers' capacity to assess against A-E ratings remain incipient even if it is expected that they shape with the future alignment of A-E ratings to the Australian Curriculum. This alignment will also improve the consistency of A-E ratings across Australia as it will lead to a national agreement on definitions for A-E ratings (see Chapter 3).

Striking the right balance between nationally-dictated policies and ability to meet local needs is a challenge

Given the current disparities of policy and practice in relation to evaluation and assessment procedures across Australia, a major challenge lies in determining what constitutes a desirable measure of consistency as against legitimate diversity. The nature of the national agenda for education is likely to be strengthened by greater consistency of evaluation and assessment procedures across jurisdictions but greater diversity offers more opportunities for innovation and adaptation to local needs. Another aspect is that the prior history, capacity and culture of evaluation and assessment across the states and territories are to be taken into account.

This tension is evident at the systemic level. States and territories have different systems of data collection – which makes comparability of data across jurisdictions more challenging – as well as distinct degrees of analytical sophistication – which potentiates innovation in some jurisdictions to the benefit of the national monitoring system. Another example concerns the diversity of approaches to senior secondary certification across states and territories.

It is clear that much of what is required in student assessment, teacher appraisal, school evaluation and system evaluation is in place in aspects of current practice across jurisdictions and school sectors. The challenge is to articulate a national strategy for each of these evaluation and assessment components which builds on the best of current practice and continues to allow flexibility of approach within agreed parameters.

There is room to improve the integration of the non-governmental sector in the overall evaluation and assessment framework

The Melbourne Declaration places strong emphasis on the fact that Australian governments "commit to working with all school sectors" on all the key areas for schooling. While through the Schools Assistance Act 2008 non-government schools have an obligation to meet national school performance and reporting requirements similar to those which apply to government schools (see earlier in this chapter), the integration of the non-government sector within the overall evaluation and assessment framework is considerably less consistent at other levels.

For instance, school evaluation practices in the Catholic and Independent systems are not mandatory and the organisation of teacher appraisal in the context of performance management processes is dissociated from state and territory School Improvement Frameworks. The typical approach for teacher and school evaluation in the non-government sector consists of giving independence to school providers to run their own procedures while state and territory authorities monitor the performance of non-governmental schools against minimum standards through registration processes (see Chapters 4 and 5). Also, the monitoring of non-government sectors is generally conducted via state or territory regulatory authorities, but reporting on their outcomes is still limited to a simple set of compliance statements and does not focus on performance. It is unclear to what extent such information for non-government schools is aggregated to the system level and analysed (see Chapter 6).

The Review Team formed the impression that there is room to improve the integration of the non-governmental sector in the overall evaluation and assessment framework. The risk of a limited integration is that there is little guarantee that evaluation and assessment procedures in the Catholic and Independent sectors are sufficiently aligned with student learning objectives and educational targets at the national and systemic levels.

Building capacity for evaluation and assessment remains a priority

The effectiveness of evaluation and assessment relies to a great extent on ensuring that both those who design and undertake evaluation activities as well as those who use their results are in possession of the proper skills and competencies. Evaluation and assessment practices in Australia benefit from outstanding expertise in areas such as standardised test development, common reporting frameworks, national comparable data on student outcomes and externally-based student assessment and relies significantly on the investigation generated by a large and active educational research community.

However, there are areas in which building capacity remains a priority. An example is the development of teacher capacity to assess against the whole range of curriculum goals to ensure consistency of A-E ratings across schools (see Chapter 3). Another area for further development, in light of the availability of rich data from student assessment and testing, is improving the data handling skills of school agents (see Chapters 3 and 5). Our interviews also revealed that the increasing complexity of some outcome reporting has not been accompanied by a good understanding by parents and other stakeholders of the concepts behind the ways the data are presented and compared. In addition, there are also considerable gaps in the development of competencies for teacher appraisal and school evaluation. There are instances of evaluators lacking credibility as they do not have specific training for their function and also concerns about the processes to select evaluators.

It is also unclear whether systematic processes are in place to identify best practices within the overall evaluation and assessment framework and ensure that they are spread and shared across educational jurisdictions and school sectors. There is a wide range of quality assurance activities developed locally within classrooms and schools, which tends not to be documented. A consequence is that the existing knowledge and information on evaluation and assessment may get lost and there is little systemic learning over time.

Policy recommendations

Establish national strategies for strengthening the linkages to classroom practice within the overall evaluation and assessment framework

Realising the full potential of the overall evaluation and assessment framework involves establishing strategies to strengthen the linkages to classroom practice, where the improvement of student learning takes place. As indicated earlier, the Review Team considers that there is no sufficient articulation of ways for the national education agenda to generate improvements in classroom practice through the assessment and evaluation procedures which are closer to the place of learning.

A major step in this direction would be a national reflection about the nature and purpose of evaluation components such as school evaluation, teacher appraisal and student formative assessment within the overall education reform strategy and the best approaches for these evaluation components to improve classroom practices. This national reflection could be promoted by national-level bodies such as MCEECDYA, ACARA and AITSL and involve representatives of educational jurisdictions, educational researchers, and other stakeholders. The final result of this reflection should be the establishment of a set of principles (or guidelines) on how to undertake or promote evaluation activities such as school evaluation, teacher appraisal, student formative assessment or the evaluation of school leadership, in ways that support national student learning objectives. The principles should build on current best practice, align with the national policy agenda and respect traditions of Australian schooling. They would communicate expectations about sound practices and suggest the development of approaches which command the confidence of all stakeholders. This national reflection would also serve as a way to share knowledge across educational jurisdictions about those evaluation and assessment procedures more closely interconnected with classroom practice.

This reflection would shed light on strategies which can contribute to reinforce the linkages between evaluation and assessment and classroom practice. Channels which are likely to reinforce such links include: an emphasis on teacher appraisal for the continuous improvement of teaching practices; ensuring teaching standards are aligned with student learning objectives; involving teachers in school evaluation, in particular through conceiving school self-evaluation as a collective process with responsibilities for teachers; ensuring that teachers are seen as the main experts not only in instructing but also in assessing their students, so teachers feel the ownership of student assessment and accept it as an integral part of teaching and learning; building teacher capacity for student formative assessment; and building teachers' ability to assess against A-E standards. It should be noted that these strategies build on teacher professionalism. Better articulating these channels within the overall evaluation and assessment framework should be part of the reflection on how best to achieve the improvement function of evaluation and assessment.

Promote greater national consistency while giving room for local diversity

The federal constitution of Australia requires that any programme of development must be established through the kind of consensus building which has characterised other aspects of the education reform agenda. The long tradition of evaluation and assessment as well as the strengths developed in this area by educational jurisdictions must be respected and built upon. However, greater consistency of evaluation and assessment practices across jurisdictions (and school sectors) would provide greater guarantees that such practices are aligned with national student learning objectives.

An important first step might be to agree protocols between educational jurisdictions and the Australian Government for the design and implementation of given evaluation and assessment procedures on the basis of the sets of principles proposed above. The protocols would involve the agreement of general principles for the operation of procedures such as school evaluation, teacher appraisal, student formative assessment or the evaluation of school leadership while allowing flexibility of approach within the agreed parameters to better meet local needs. The approach could follow the example of the Principles and Protocols for Reporting on Schooling in Australia (MCEECDYA, 2009). For each of the evaluation components on which principles and a protocol would be agreed, a number of fundamental issues should be addressed, including: how to combine the accountability and improvement functions; the scope in relation to the national agenda; aspects to be assessed; reference standards; the role and nature of externality; and the extent of transparency.

The protocols should come along with clear goals, a range of tools and guidelines for implementation. They should permit better consistency of evaluation practices across educational jurisdictions while leaving sufficient room for local adaptation. This could imply requiring educational jurisdictions to develop action plans at the local level aligned with national protocols. The goals defined at the national and the jurisdiction level should be complementary in order to avoid conflicting messages to schools.

Improve the integration of the non-governmental sector in the overall evaluation and assessment framework

Evaluation and assessment practices in the Catholic and Independent sectors are very diverse and, with the exception of the reporting requirements which apply to all schools across Australia, display limited alignment with those in place in state and territory schools. As a result, in spite of well-consolidated practices in the non-government sector, there is limited guarantee that those practices are aligned with the national education agenda. This is the case in spite of the high degree of collaboration among the government and non-government sectors in many states and territories.

Regarding evaluation and assessment procedures closer to the classroom (*e.g.* school evaluation, teacher appraisal), a possible solution to better integrate the non-governmental sector in the overall evaluation and assessment framework is for the non-government sector to be part of the Protocol agreements suggested above to reach greater national consistency towards the national education agenda. This could become another requirement for non-government schools to receive public funding in a way similar to the reporting requirements. Certification of the adherence to the protocols could then be part of registration processes whereby educational state and territory authorities grant authorisation for non-governmental schools to operate.

At the system level, and in order to monitor the performance of non-government schools, non-government schools could be compelled to adhere to state and territory administrative data collections and be part of common performance summary reports for schools in all sectors. Consideration could also be given to extending the mandate of state and territory Auditor General Offices to the review of all schools receiving government funding (see Chapter 6).

Further develop some articulations within the overall evaluation and assessment framework

The process of developing an effective evaluation and assessment framework should give due attention to: achieving proper articulation between the different evaluation components (*e.g.* school evaluation and teacher appraisal); warranting the several elements within an evaluation component are sufficiently linked (*e.g.* teaching standards and teacher appraisal); and ensuring processes are in place to guarantee the consistent application of evaluation and assessment procedures (*e.g.* consistency of teachers' A-E ratings).

For example, as explained in the previous section, there is room to better define the articulations between: teacher appraisal and student assessment (see Chapter 4); school evaluation and system evaluation (see Chapter 5); and school evaluation and teacher appraisal. Regarding the latter articulation, analysis from TALIS (OECD, 2009) suggests that school evaluations can be an essential component of an evaluative framework which can foster and potentially shape teacher appraisal and feedback. Given that the systems of school evaluation and teacher appraisal and feedback have both the objective of maintaining standards and improving student performance, there are likely to be great benefits from the synergies between school evaluation and teacher appraisal. To achieve the greatest impact, the focus of school evaluation should either be linked to or have an effect on the focus of teacher appraisal (OECD, 2009). This indicates that school evaluation should comprise the monitoring of the quality of teaching and learning, possibly include the external validation of school-based processes for teacher appraisal (holding the school director accountable as necessary), and school development processes should explore links to the evaluation of teaching practice (see Chapters 4 and 5). In the context of school self-evaluation, it is also important to ensure the centrality of the evaluation of teaching quality and the feedback to individual teachers.

Examples of linkages within single evaluation components which need to be reinforced include the association between teacher appraisal and teacher professional development (see Chapter 4), the alignment of teaching standards with student learning objectives (being currently achieved by the work of AITSL), the articulation between school self-evaluation and external school evaluation (see Chapter 5), the relationship between standardised student testing and student formative assessment (see Chapter 3) and the alignment between A-E ratings and the Australian Curriculum (being currently addressed in work led by ACARA).

Finally, moderation processes are vital to ensure the consistency of the application of evaluation and assessment procedures and in this respect priority should be given to teachers' A-E ratings and the appraisal of teachers against teaching standards.

Sustain efforts to improve capacity for evaluation and assessment

The development of an effective evaluation and assessment framework involves considerable investment in developing competencies and skills for evaluation and assessment at all levels. This is even more the case in the context of a national education reform of considerable dimensions which places particular emphasis on national assessment, reporting requirements, the Australian Curriculum, teaching standards, and capacity to assess against A-E standards. The reform has advanced on solid grounds at an impressive pace but, understandably, the time elapsed since it got underway has not yet allowed building the levels of capacity for evaluation and assessment necessary to realise the full potential of the overall evaluation and assessment framework.

As a result, it is clear that an area of policy priority is sustaining efforts to improve the capacity for evaluation and assessment. Areas in which the Review Team believes considerable investment should be made are: developing teachers' capacity to assess against A-E standards; improving the skills of teachers for formative assessment; improving the data handling skills of school agents; and facilitating the understanding by parents and other stakeholders of the concepts behind the ways the data are presented and compared. Capacity building through adequate provision of initial teacher education and professional development should be a priority making sure provision is well aligned with the national education agenda. Other strategies involve the provision of support materials; scoring guides and exemplars of different A-E ratings; and website platforms proposing formative teaching and learning strategies. (See Chapters 3 and 5).

Another area which deserves attention relates to skills and competencies for teacher and school evaluation. The Review Team formed the impression that these are uneven across educational jurisdictions and schools. A more systematic approach to training for teacher and school evaluation within educational jurisdictions should be considered. Since evaluation has strong stakes for the units assessed and since school outcomes heavily depend on individual relations and co-operation at the school level, successful feedback mechanisms require particular attention to developing competencies and defining responsibilities in the evaluation process. In addition, competencies for using feedback to improve practice are also vital to ensure that evaluation and assessment procedures are effective. Assessment for improvement requires the inclusion of actors such as teachers in the process of school development and improvement. As a result, for instance, it is pertinent to include training for evaluation in initial teacher education alongside the development of research skills. Similarly, the preparation to become a school leader is expected to include educational leadership with some emphasis on feedback mechanisms. (See Chapters 4 and 5).

Another area to explore is building capacity at systemic (jurisdiction) level to ensure an effective use of the results generated by evaluation and assessment activities. Ontario presents an interesting example of a focused body within the central department to promote and build capacity throughout the education system (in this case to improve literacy and numeracy) and early results indicate a positive impact (Box 2.3). This draws in part on an information system that allows the monitoring of the impact of particular initiatives introduced by the education department.

Box 2.3 Strengthening monitoring on literacy and numeracy in Ontario

In 2004, the Ontario Ministry of Education launched a Literacy and Numeracy Strategy, part of which included the creation of a Literacy and Numeracy Secretariat (LNS). The aim of the LNS is to improve student outcomes by building instructional and leadership capacity throughout the education system. An evaluation of the impact of initiatives introduced by the LNS concludes that they have had "a major, and primarily highly positive, impact on Ontario's education system" (Canadian Language and Literacy Research Network, 2009). The evaluation was conducted over a two-year period and drew on much information including surveys to school principals, teachers and LNS Student Achievement Officers in public and Catholic schools in both the English and French systems. The report presents evidence that the LNS has worked effectively with the Education Department and educators to build capacity and improve student outcomes and to positively impact school boards and schools. There have been sustained improvements on key measures of student performance in reading, writing and mathematics in Grades 3 and 6 (although improvements are smaller in mathematics) and these are observed for key student groups such as English Second Language learners and special needs students.

The report notes the expansion of research, evaluation, planning and data management capacity at the Ministry and school board levels and evaluates that this has led to a better understanding of how to address challenges and how to learn from successes. Although this expansion has focused on improving literacy and numeracy skills, the report argues that the heightened use of evidence, research, evaluation and data throughout the system should provide long-term benefits in other areas.

The report commends the commitment to data-based decision making at both the system and local levels. One major initiative is a new information system "Ontario Statistical Neighbours (OSN)" which allows school-level analysis of performance, context and school programmes/interventions. All publicly funded schools are included in the information system (Ontario Ministry of Education, 2007). Indicators are drawn from Statistics Canada and Ministry's information systems and include performance indicators from Ontario assessments, demographic indicators on proportion of students from low income or university educated families and second language learners at school, data on school characteristics such as location, size and proportion of students with special education needs or instructional language support needs, and data on LNS programmes and special interventions. OSN can provide useful information to the system, school boards and schools and offers four analytical modules:

- Information Centre – full analysis on school demographics, socio-economic characteristics, school programme information and school performance indicators, at the system, regional or school board level.
- Schools Like Ours – analysis of similar schools on school demographics, school programme information and school performance indicators, at the system, regional or school board level.
- Geographic – analysis of school performance by geographic location, *e.g.* cities or towns.
- Performance – analysis of school performance on individual assessment areas or across assessment areas, at the system level.

The report comments on the potential for OSN "to enable quick and accurate identification, monitoring, and intervention with schools and groups of schools" and notes that it had already provided information to all but two school boards, mostly queries on schools with particular challenges or the identification of similar schools. The report notes the importance of promoting awareness among school boards of the full analytical potential of OSN.

Sources: Canadian Language and Literacy Research Network (2009) and Ontario Ministry of Education (2007).

Finally, at the national level, there is a need to put in place systematic processes to identify best practices within the overall evaluation and assessment framework and ensure their dissemination across educational jurisdictions and school sectors. This objective can greatly benefit from the operation of national-level institutions such as ACACA, ACARA and AITSL.

Notes

1. It should be noted that Australia does not have one single framework that was designed as a whole but instead has a series of components operating at different levels that will be referred to as the "overall framework" throughout this report.
2. The Productivity Commission is the Australian Government's independent research and advisory body on a range of economic, social and environmental issues affecting the welfare of Australians. Its role, expressed simply, is to help governments make better policies in the long-term interest of the Australian community (*www.pc.gov.au*).

References

Australian Government (2010), *Country Background Report for Australia*, prepared for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, Department of Education, Employment and Workplace Relations, Canberra, available from *www.oecd.org/edu/evaluationpolicy*.

Bortoli, L. de, and S. Thomson (2010), *Contextual Factors that Influence the Achievement of Australia's Indigenous Students: Results from PISA 2000–2006*, Australian Council for Education Research, Camberwell, Victoria, *www.acer.edu.au/documents/pisa-indigenous-contextual-factors.pdf.*

Canadian Language and Literacy Research Network (2009), *Evaluation Report: The Impact of the Literacy and Numeracy Secretariat: Changes in Ontario's Education System*, Canadian Language and Literacy Research Network, *www.edu.gov.on.ca/eng/document/reports/OME_Report09_EN.pdf.*

Cresswell, J. and C. Underwood (2004), "Location, Location, Location: Implications of Geographic Situation on Australian Student Performance in PISA 2000", *ACER Research Monograph No. 58*, ACER, Camberwell, *www.acer.edu.au/documents/PISA_RM58PISALocation.pdf.*

Department of Education and Early Childhood Development (2010), *Department of Education and Early Childhood Development – Annual Report 2009-10*, State Government Victoria, Melbourne, *www.eduweb.vic.gov.au/edulibrary/public/govrel/reports/200910deecdannualreport.pdf.*

Grattan Institute (2010), *What Teachers Want: Better Teacher Management*, Melbourne, available from *www.grattan.edu.au*.

MCEECDYA (2009), *Principles and Protocols for Reporting on Schooling in Australia*, Ministerial Council for Education, Early Childhood Development and Youth Affairs, *www.mceecdya.edu.au/mceecdya/nap_principles__protocols_for_rep_on_school_2009,27896.html*.

MCEETYA (2008), *Melbourne Declaration on Educational Goals for Young Australians*, Ministerial Council on Education, Employment, Training and Youth Affairs, Canberra.

OECD (2009), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, OECD, Paris.

Ontario Ministry of Education (2007), *Ontario Statistical Neighbours – Informing our Strategy to Improve Student Achievement*, The Literacy and Numeracy Secretariat, Ontario Ministry of Education, *www.edu.gov.on.ca/eng/literacynumeracy/osneng.pdf.*

Productivity Commission (2010), *Strengthening Evidence-based Policy in the Australian Federation*, Roundtable Proceedings, Canberra, 17-18 August 2009.

Steering Committee for the Review of Government Service Provision (2009), *Overcoming Indigenous Disadvantage: Key Indicators 2009*, Productivity Commission, Canberra, *www.pc.gov.au/__data/assets/pdf_file/0003/90129/key-indicators-2009.pdf.*

Chapter 3

Student assessment

Student performance in Australia is assessed by a wide range of instruments, ranging from national standardised tests to ongoing daily formative assessment in the classroom. At the national level, both full-cohort and national sample assessments of Australian students are conducted, the results from which are used as key performance measures towards national goals. At the system level, many state and territory governments administer testing with both diagnostic and monitoring purposes. States and territories are also responsible for externally-based summative assessment, in particular in view of assessing students for secondary education certification. At the school level, student assessment plays the key role in informing schools and teachers about students' individual achievement through teacher-based summative and formative assessments. A major asset is that a coherent framework for the assessment of student learning is in place in Australia. Other strengths include the credibility of NAPLAN results among school agents; the moderation processes and dedicated tools to support student assessment; the existence of consolidated assessment practices for secondary school qualifications; good practices of formative assessment; and the reliance on teacher-based summative assessment. Priorities for future policy development include developing national consistency while respecting state and territory assessment strengths and cultures; reinforcing the assessment validity of NAPLAN; establishing safeguards against overemphasis on NAPLAN; strengthening teachers' capacity to assess student performance against the Australian Curriculum; building teachers' competence to use student assessment data; maintaining the centrality of teacher-based assessment; and increasing the visibility of the Australian Government's goals for formative assessment.

This chapter focuses on approaches to student assessment within the Australian overall evaluation and assessment framework. Student assessment refers to processes in which evidence of learning is collected in a planned and systematic way in order to make a judgement about student learning (EPPI, 2002). This chapter looks at both summative assessment (assessment *of* learning) and formative assessment (assessment *for* learning) of students.

Context and features

An extensive system of student assessment

Australia has an extensive system of national-, system- and school-level student assessment. Student performance in Australia is assessed by a wide range of instruments, ranging from national standardised tests to ongoing daily formative assessment in the classroom. At the national level, ACARA is responsible for both full-cohort and national sample assessments of Australian students, the results from which are used as key performance measures towards national goals. At the system level, many state and territory governments administer testing such as senior secondary testing, which is used together with the national assessments as part of system evaluation. At the school level, student assessment plays the key role in informing schools and teachers about students' individual achievement through teacher-based summative and formative assessments. This is detailed below.

National Assessment Program in Literacy and Numeracy

The annual National Assessment Program in Literacy and Numeracy (NAPLAN) commenced in 2008 and is administered to all students in Years 3, 5, 7 and 9, in both the government and non-government school sectors. It replaced the eight state- and territory-based assessments that were used previously. Results are reported relative to five national achievement scales: reading, writing, spelling, grammar and punctuation, and numeracy. Developed in advance of the Australian Curriculum, NAPLAN, according to its developers, reflects content and skill clusters that generally were common across the various state and territory assessments administered prior to NAPLAN. NAPLAN tests are developed collaboratively by the states and territories, the Australian Government, and representatives of the Catholic and Independent school sectors.

NAPLAN presents items reflecting increasingly challenging understanding and skill in each cluster at succeeding year levels and uses a consistent 1-10 national achievement scale to classifying student performance. For example, students in Year 5 typically score in bands 3-8, those in Year 9 typically fall in bands 5-10. The bands are used to define what constitutes the national minimum standard for each year level, which indicates that the student has demonstrated the basic skills of literacy and numeracy expected in that particular year level. Results at a lower level band indicates that the student has not achieved the basic skills expected for that year level and may need focused intervention and additional support to help achieve these skills. For example, a score in band 2 indicates the national minimum standard for students in Year 3; those scoring in band 1 need help and those scoring in band 3 and above are considered above the national minimum standard. Similarly, national minimum standards are established for Years 5, 7, and 9: bands 4, 5, and 6 respectively.

Schools are provided with a detailed report on their (individual) students' results. Information on the full range of NAPLAN achievement is provided, including the number of students in each band at each year level. Schools are expected to use this information to monitor student progress and to identify students in need of additional support. Tools and strategies to support teachers' and schools' analysis and use of the data vary across the states and territories. Victoria, for example, provides its schools with annual Data Service Reports, which summarise overall and subgroup results by grade and achievement scale in comparison to the state and nation, item analyses, trend data and individual student performance, in addition to specific guidance on how results should (and should not) be interpreted. New South Wales has developed a sophisticated School Measurement, Assessment and Reporting Toolkit (SMART), which enables teachers and schools to analyse the data themselves at the level of individual student, classroom and school (see Box 2.2, Chapter 2).

Individual student reports are prepared and distributed to parents. The report shows each child's results in comparison with all other children in Australia at the same year level who took the test. The displays, for each of the five achievement scales, typically show:

- National average;
- National minimum standard;
- Range for the middle 60% of students nationally;
- Indication of whether the child has achieved the national minimum standard;
- School average for the same year level (for some states and territories);
- The items the student successfully responded to and those s/he did not (for some states and territories).

That students' scores are relative to a consistent 10-band national achievement scale covering all years means that parents can directly compare their children's scores from one year to those on subsequent NAPLAN assessment. The back page of the report provides a table that briefly describes the skills that students scoring at each band have typically demonstrated.

The results of NAPLAN are published at the school level on the *My School* website (see Chapter 5) and are also extensively used for monitoring performance at jurisdiction and national levels, whereby informing policy deliberations at both levels (see Chapter 6).

Other elements of the National Assessment Program

The Australian National Assessment Program (NAP), in addition to NAPLAN, also includes triennial tests in information and communication technology literacy in Years 6 and 10, science literacy in Year 6, and civics and citizenship in Years 6 and 10. These assessments are designed primarily to monitor national and jurisdictional progress; however participating schools receive their own students' results and the school's results. These can provide useful information to classroom teachers and assist with curriculum planning (see Chapter 6).

As part of NAP, Australian students also participate in international tests, including the Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). The objective is to benchmark

Australian education performance internationally by monitoring student progress over time against international standards (see Chapter 6).

Student assessment practices in states and territories

Each state and territory has control over its student assessment system. Approaches typically involve a mix of summative and formative assessments and a combination of externally-based and teacher-based assessments. The progress of students is assessed against specific curriculum standards for each state and territory defining what students should know, understand and be able to do.

Externally-based assessment

All states and territories have forms of system-wide testing in schools (see Chapter 6). Table 17 in the Country Background Report (Australian Government, 2010) contains links to the state and territory agencies and authorities that develop, administer and co-ordinate many system-wide assessments. They are responsible for delivering services, such as testing and moderation between schools and sectors. Examples of system-wide tests are:

- The Queensland Comparable Assessment Tasks (QCATs): These are performance-based assessment tasks administered each year to students at both primary and secondary level in Years 4, 6 and 9 to test directly against the curriculum. Their purpose is to provide information about student learning in targeted learning areas of English, mathematics and science and to help promote consistency of teacher judgments across the state.
- The Western Australian Monitoring Standards in Education (WAMSE): Standardised assessments in science and society and environment with no stakes for students. In 2010 all Year 5, 7 and 9 students from government and Catholic schools and many in Independent schools were assessed in these two learning areas, providing valuable information for Western Australian systems, schools, teachers, parents and students. This complements the information from NAPLAN.
- The New South Wales Essential Secondary Science Assessment (ESSA): Mandatory test for students in government schools who have completed two years of secondary schooling and learning in science (optional for students in non-governmental schools). It is a diagnostic test which is used to support teaching and learning.

In some states and territories, Year 10 students sit a formal test developed at the state and territory level around the curriculum goals and standards towards gaining a record of achievement, such as a School Certificate, or Year 10 Certificate.

Teacher-based summative assessment

Teachers' professional judgment historically has been integral to assessment in Australia and the primary means, particularly in early and middle school years, through which students and their parents have been apprised of student progress. Assessment tools chosen by teachers vary to ensure the students are provided with ample opportunity to demonstrate their ability against the curriculum across a range of contexts. For example, assessment of student performance in the classroom may be done both formally, through

tests, short and long constructed response tasks, projects or rich tasks requiring the planning, development and presentation to peers, and informally through observations and discussion.

Common assessment methods in primary school settings which are used to provide evidence of achievement using outcomes include observing and recording student achievement as it occurs, mapping progress through the collection of student work samples over a period of time, tasks that incorporate the application of understandings and learning processes in a set project and analysis of non-print-based work samples in areas such as the visual arts.

Secondary school-level assessment practices are varied, ranging from laboratory experiments, essay writing, research papers, presentations, demonstrations, projects, assignments, tests and school-based examinations.[1] Schools have the responsibility for determining assessments that best suit the students, including qualitative and quantitative assessment. There are some school-designed, year level assessments. These assessments are often referred to as "common tests" and are generally focused on students in the middle years of schooling. They are usually designed and implemented by schools in an effort to identify student achievement groups and plan for the allocation of learning support.

Collaborative moderation is a key strategy in validating consistency of teacher judgement and marking and it typically occurs within schools, between schools and across sectors.

Formative assessment

Formative assessment is an integral component of student assessment for Australian teachers. In contrast to summative assessments *of* learning that are mandated top-down, assessments *for* learning, so-called formative assessments, occur bottom-up, within the actual context of classroom teaching and learning. The last decade has witnessed an explosion of worldwide interest in formative assessment, fuelled in large part by Black and Wiliam's landmark meta-analysis showing the strong effects of formative assessment on student learning, particularly for low ability students (Black and Wiliam, 1998; OECD, 2005).

The use of data is key to the idea: to be considered formative, assessment evidence must be *acted upon* during the course of classroom instruction. Rather than focusing backward on what has been learned, formative assessment helps to chart the learning road forward, by identifying and providing information to fill any gaps between the learners' current status and goals for learning (Sadler, 1989). Assessment is used to elicit students' understanding in order to provide immediate feedback to teachers and students that can be used "to form" subsequent teaching and learning (Wiliam and Thompson, 2007). In some recent formulations, the involvement of students in the process raises a third function of assessment: assessment as learning, which focuses on students reflecting on and monitoring their own progress to inform their future learning goals (as reflected in the Melbourne Declaration, MCEETYA, 2008).

The use of formative assessment is common in Australian schools. In general, teachers and students are clear on learning expectations and work collectively to achieve them. Teachers communicate expected learning goals to students and provide ongoing feedback to help students attain the goals. The day-to-day teaching and learning process typically includes activities such as classroom interactions, questioning, immediate

feedback, guidance on how to close learning gaps and student engagement in self- and peer- assessment.

Summative assessment in senior secondary education

All states and territories also have some form of senior secondary completion assessment covering both the government and non-government sectors. Years 11 and 12 subject-based exams and vocational education and training exams figure heavily in the quality of Australia's education system and may be the most important assessments for students, as they are used for admission to tertiary education, work placement and employment. While these exams vary by state and territory, they provide the basis for the Australian Tertiary Admission Rank (ATAR), which combines students' relative performance with moderation procedures to place students on a common scale across all locales (except Queensland). While these exams vary by state and territory, there are moderation processes in place across Australia. Students are given a wide choice of subjects, which essentially constitute the core of their secondary education. While each state and territory has its own system and own set of procedures for developing and approving courses (including specified learning goals, content, and exams), most combine student performance on external exams at the end of Year 12 with moderated, teacher judgments of coursework performance to arrive at scores for senior secondary certificates and high school completion. External exams are derived from a combination of multiple-choice, short answer and extended response tasks. Course work includes a variety of tasks, including extended performances. Examination systems in the Australian Capital Territory and Queensland are more school-determined and based, but achievement standards and scoring are externally moderated.

In Queensland, the moderation processes for the Senior Certificate (Year 12) involve subject-based panels of expert teachers providing advice to schools on the quality of their assessment programme and their judgments of quality of student performance based on sample portfolios. The system involves follow-up where panels identify difficulties. There is negotiation of the final results to be recorded on the Senior Certificate (Sebba and Maxwell, 2005).

A-E Reporting System

In response to parental feedback at national, state and territory forums suggesting that parents were confused by the different reporting scales and mechanisms used across schools, the Australian Government brought a degree of standardisation to teachers' judgments by requiring in 2005 that each state and territory adopt a common five-point scale as a condition for federal funding. At each year level from Year 1 to Year 10, teachers have to report students' achievements to parents using an A-E (or equivalent) framework. Defined specifically by each state and territory, generally the points on the A-E scale represent:

- A means well above standard;
- B means above standard;
- C means student at expected standard at time of report, on track to proficiency;
- D means below standard;
- E means well below standard.

A-E ratings are intended to assess the full range of learning expected of students. States and territories have developed guidelines and definitions for each of the A-E levels, variously labelled as letters (A-E) or descriptive categories (*e.g.* advanced – developing). States and territories vary in the specificity of the definitions and guidance they provide to support consistent judgments across teachers and schools. For example, Victoria provides teachers with detailed standards (*Victorian Essential Learning Standards*) co-ordinated with expected progression points, assessment maps and assessment modules to gauge student progress. Reporting software enables teachers to enter assessment scores and other ratings for components of each standard and the system automatically aggregates these scores into overall ratings for each student. To support consistency, Victoria also examines the relationship between the distribution of students' A-E ratings and NAPLAN results.

There is currently a proposal to link A-E standards to the Australian Curriculum. The work, led by ACARA, will start in 2011 on a common approach to the achievement standards across states and territories including trialling and validation. In future, part of the work to align A-E with the Australian Curriculum will involve national agreement on definitions.

Box 3.1 lays out validity criteria by which individual student assessments may be judged. These criteria were used by the Review Team in the subsequent analysis of strengths and challenges to student assessment in Australia.

Box 3.1 Validity criteria to judge individual student assessments

Validity resides in evidence of the extent to which an assessment embodies the characteristics that support its intended purpose(s) and the extent to which the scores from an assessment yield meaningful inferences to support intended decisions and uses. This suggests two basic kinds of criteria for judging the quality of an individual student assessment: one set based on the attributes or the characteristics of the assessment/test itself, and the other based on the validity of score interpretations. Assessments must not only yield technically sound measures of student learning, but also provide results that are appropriate for intended uses. For example, a reading test may provide an accurate measure of students' reading skills, but not provide a reliable diagnosis of the source of students' reading difficulties. Such a test could be appropriate for purposes of evaluation or progress monitoring, but not appropriate for formative purposes.

The criteria below originally were developed based on core concepts in the United States' *Standards for Psychological and Educational Testing* (AERA, APA and NCME, 1999) (See also Linn *et al.*, 1991; Herman, 2010). The criteria are consistent with MCEECDYA's Principles and Protocols for Reporting on Schooling in Australia (MCEECDYA, 2009).

Criteria related to assessment characteristics:

- *Learning-based, aligned with standards*
  - Aligned with significant learning goals
  - Comprehensive in representation of target constructs/domains intended for assessment, in both content and cognitive demands
  - Linked with expected trajectories of learning
- *Fairness*
  - Accessible, enabling all students to show what they know
  - To the extent possible, free of knowledge and skills that are irrelevant to the target of the assessment (*e.g.* language demands)
  - Sensitive to a range of student abilities and learning status; appropriate for students at the range of developmental levels likely in assessed population
- *Utility*
  - Timely
  - Useable/interpretable by teachers and/or students; provides actionable feedback for intended users
  - Instructionally useful, at the right grain size to guide subsequent, intended decision-making and action
- *Consequences:* models sound pedagogy and supports professional practice
- *Credibility:* educators; students and parents; public
- *Feasibility*: time cost of, and capacities needed for administration, scoring, reporting and use

Criteria related to validity of score interpretations:

- *Technical soundness*
  - Score reliability at level of intended use(s) (*e.g.* if assessment is formative, scores provide reliable diagnostic information)
  - Reliability of scoring
- *Generalisability and extent of transfer*
- *Instructional sensitivity* (*i.e.* that test scores reflect the quality of instruction)
- *Fairness/lack of bias*
- *Comparable*, as necessary: across sites; across time, within and across years

In applying these criteria to individual assessments, researchers have noted the need for system validity. That is, jurisdictions rarely use a single measure for purposes of accountability and improvement – and indeed they should not – but instead draw on a variety of measures that can be considered an assessment system. Standards for judging such systems have been advanced (Baker *et al.*, 2002; Herman, 2010).

Strengths

A coherent framework for the assessment of student learning is in place

A range of provisions for the assessment of student learning are established, including: NAPLAN; triennial sampled-based assessments of ICT literacy, science literacy, and civics and citizenship; international assessments (*e.g.* PISA and TIMSS); A-E ratings; and senior secondary certificates and vocational education and training certificates. This set of assessments results in a coherent system of assessments of learning that potentially can provide a comprehensive picture of student performance relative to Australia's goals for student learning. That is, while NAPLAN and other periodic assessments provide a national barometer of performance necessarily on a limited set of standards (*i.e.* those that can be measured within limited testing time), A-E reporting requirements and secondary certificates provide a structure for linking accountability to a fuller set of national and/or educational jurisdictions' expectations for student learning. Performance on international measures enables policy makers and the public to monitor student progress over time against that in other countries and at the same time serves to document the extent to which Australia's standards are consistent with leading countries in the world; Australia's children are internationally competitive; and Australia's assessments provide valid evidence relative to world-class performance.

The set of assessments also provides a structure for potentially integrating accountability and instructional improvement from early education through secondary completion. A-E reporting requirements link ongoing classroom instruction and grading with existing standards and offer a mechanism for linking individual student accountability, classroom instruction and wider accountability. For example, the Victorian Registration and Qualifications Authority (VRQA) provides performance reports on every government school that include students' A-E performance as well as NAPLAN performance and other indicators of school climate and satisfaction. The school's A-E summary scores are simply an aggregation of the individual student report card grades and standards that are the focus of classroom instruction and reported individually to parents. Giving high profile to multiple indicators, the VRQA school reports presumably also may counteract over-reliance on any one of them, *e.g.* NAPLAN, that otherwise could potentially narrow the curriculum.

Moreover, as evident during discussions with teachers in a number of jurisdictions, the A-E reporting requirements not only encourage coherence of classroom instruction with national and/or states and territories standards, but also foster coherence and consistency in interpretation within and across schools. Conversations with teachers in several primary schools suggested that they frequently discussed students' work related to A-E standards and strategies for improving student progress toward the standards. Secondary school certificates provide a similar structure for promoting coherence in standards and learning goals within and across schools and for linking individual student and school accountability.

NAPLAN results are credible among school agents and deemed useful

Most stakeholders find NAPLAN results a credible source of evidence. While a number of stakeholders we talked with across the states and territories were concerned about an over-reliance on NAPLAN and its basic skills focus (which we discuss further below), they found the results trustworthy. It is recognised that NAPLAN enables greater

consistency, comparability and transferability of results across jurisdictions in a way that was not possible under the previous jurisdiction-based testing system. The use of a common scale is also valued as it provides significant information about the performance of, and growth in, individual student achievement which can be monitored over time and add a longitudinal dimension to the data.

In multiple jurisdictions teachers spoke of the diagnostic value of NAPLAN results in helping them to identify students who had not reached expected standards and in helping them to identify specific weaknesses. In light of the absence of subscale reports it is interesting that teachers and principals with whom we spoke characterised the NAPLAN tests as diagnostic. That is, formative assessment theorists suggest that diagnostic assessments reveal the source of student difficulties, while NAPLAN results provide reliable diagnostic data only on student strengths and weaknesses relative to the various achievement scales (reading, writing, grammar and punctuation, spelling and numeracy). For example, teachers mentioned using the individual questions that students answer correctly and incorrectly to diagnose student needs, yet, in isolation, information based on a single item is not reliable. Teachers and principals also seem to find utility in the suggestive analyses of what knowledge and skills are typical of students scoring at particular achievement levels. The sophisticated software and standard reports provided in some states and territories – for example, Victoria and New South Wales – enable and encourage teachers to dig deeper into the specifics of item performance and they appear to value such detailed analyses.

There is evidence that NAPLAN is technically sound

The trust placed in NAPLAN findings seems well justified from the perspective of the reliability and precision of reported scores. The *NAPLAN 2008 Central Analysis Technical Report* (ACER, not dated) documents the technical quality of the assessments through reliability, discrimination, item fit, and differential item functioning indices which generally suggest that the measures are technically sound for non-Indigenous students. Vertical and year-to-year equating and scales are well and carefully constructed, as are proficiency estimates and relevant cut off points. The authors of the Technical Report clearly prize the psychometric soundness of the reported NAPLAN findings and take care to assure the reliability of each reported score. For example, the 2008 Technical Report used dimensionality analyses to examine the feasibility of reporting a combined literacy scale, composed of student performance on the reading, writing, grammar and punctuation, and spelling scales, and of reporting subscales for numeracy. In both cases they found that the psychometric analyses did not support any change in reporting. At the same time, the timing and speed with which NAPLAN was accomplished has also given primacy to psychometric indices of quality rather than broader validity criteria such as alignment with significant learning goals, comprehensiveness in representing target constructs, or instructional sensitivity (see Box 3.1).

It is clear that NAPLAN draws upon good expertise in designing and reviewing the test, excellent research knowledge and technical expertise in developing the achievement scale and world-class psychometric methods in analysing and reporting the results in a meaningful way for teachers and parents.

Summative student assessment is adequately supported by moderation processes and dedicated tools

The tools and resources developed by educational jurisdictions to support their schools and teachers' use of the A-E reporting scales, such as assessment tools and measurement standards linked to school curricula, appear very valuable to teachers. Typical support includes the use of software for A-E reporting and NAPLAN data analysis software. Direct school specific support is typically provided in the use of senior secondary performance data to inform instructional strategies at the classroom level – including the development of effective assessment instruments. (As we note below, there might be a case for the Australian Government to share good practices in this area across educational jurisdictions).

Similarly, procedures adopted by educational jurisdictions and particular schools for moderating A-E judgments and senior secondary assessments also are models for increasing the utility and consequences of assessment: not only does the moderation process facilitate common understanding of year level proficiency standards, how students' learning toward them develops and potential obstacles to progress, but also fosters the development of professional learning communities that can provide crucial support for improving opportunities for student learning and building teacher capacity (McLaughlin and Talbert, 2006).

Assessment for secondary school qualifications is well established

The diversity of approaches to senior secondary certification across states and territories can be seen as strength. Each system offers students' choice in secondary pathways and coursework to meet individual interests. Each has in place review processes and procedures to assure rigorous course syllabus and standards and assessments for certifying student accomplishment. Each has in place strong moderation procedures to assuring consistency and comparability of qualifications within and across schools and subjects. Based on site visits to selected schools in New South Wales, Queensland, Victoria and Western Australia, examination systems function to clearly communicate goals and expectations for students to teachers and students alike. Students appear to understand the criteria by which they will be judged and get ongoing feedback to support their progress and success. There appear to be strong cultures – from the perspectives of teachers, students and parents – supporting the unique systems in each jurisdiction. The Australasian Curriculum, Assessment and Certification Authorities (ACACA), as a body, provides a means for monitoring and enhancing developments in senior secondary curriculum and certification. While the operations of the Vocational Education and Training sector were less visible in our visit (given that the bulk of VET courses are undertaken in specific institutions not in the scope of the Review), Australia's efforts to assure choice, rigour and relevant job training to prepare students for work post high school completion seems to be effective.

Good practices of formative assessment

While the Review Team did not have the opportunity to observe any classes, both teachers and students provided indications that good practices of formative assessment are established. Students spoke of the targeted and frequent feedback they received to help them reach established learning goals. Teachers seem to communicate learning goals to students, engage students in the learning process, and use data from the learning process to inform subsequent instruction. Teachers also appear to engage students in self-assessment

and peer-assessment. Students, particularly at the secondary level, are given responsibility for their own learning. Feedback from ongoing work and assignments shows them where they are relative to expected goals and they are expected to act to close any gaps.

Similarly, teachers and school principals spoke of their strong commitment to the use of data to improve student learning and their own accountability for student learning. Teachers appear committed to using ongoing classroom data. Throughout school visits, we heard extended examples of teachers' formative use of data to identify individual students' strengths and weaknesses and to take appropriate steps to promote subsequent progress. This is supported by tools available to identify students' strengths and weaknesses. For instance, schools in Victoria have access to a range of on-demand diagnostic assessments through the Department of Education and Early Childhood Development, the Victorian Curriculum and Assessment Authority, and third parties such as the Australian Council for Educational Research. A specific example is the English Online Interview (EOL), the only mandated standardised assessment in Victoria in addition to NAPLAN, which is a diagnostic/instructional tool for teachers, to inform planning and personalised teaching strategies (results are not reported to students or parents).

Existence of diagnostic tests upon commencement of primary education

A positive development is the establishment of diagnostic tests for students upon commencement of primary education in most states and territories and non-government sector schools to determine their educational and skill level when they first enter school. Examples of such diagnostic assessments are the School Entry Assessment in South Australia, the Performance Indicators in Primary Schools assessment tool in Tasmania and the Australia Capital Territory, the Best Start Assessment in New South Wales, and the Online Interviews in English and Mathematics in Victoria and Western Australia. These diagnostic tests bring considerable support to teachers through: gathering information about the knowledge, skills and understanding children first bring to school; recording the developmental stage the child is in; and using the information to plan learning programmes to meet the needs of the child.

There is considerable reliance on teacher-based summative assessment

There is a good focus on covering a broad range of evidence on student performance through teacher-based assessment in overall student summative assessments.[2] Teachers' continuous classroom-based assessments are included in students' grades and typically contribute to the school-leaving certificate report. Although there are challenges about the unevenness of teacher grading both within and between schools, the practice of giving considerable weight to teacher-based assessment in student summative assessments is nevertheless important. Teachers have many more opportunities to observe students over time in performing a variety of tasks, and in this sense their observations are more comprehensive than a single, high stakes assessment can ever be. Due to its continuous nature, teacher-based assessment often allows for important achievements to be measured that could not be captured in an external examination, such as extended projects, practical assignments or oral work. Teachers are also less likely to "teach to the test" when they are able to take into consideration a range of experiences and observations of student performance.

Our site visits revealed rich examples of teachers' collaboration to analyse student work relative to expected standards, to discuss learning issues, and to plan next steps for

instruction, including feedback to students. Teachers described how such collaboration enabled them to reach consensus on A-E expectations for student learning and to draw on each others' expertise to support their students' progress. Individual teachers and students also described how assessment for learning was working for them and the value of being clear on what was expected and getting ongoing feedback about progress. At the same time, such assessment practices appeared quite variable over the schools the Review Team visited, and most teachers struggled with how to respond when their initial – or even multiple – rounds of teaching and learning failed to get students to expected learning goals.

Challenges

NAPLAN has certain limitations in its alignment with student learning objectives

NAPLAN was developed and implemented prior to the introduction of the Australian Curriculum and thus may not be closely aligned with it. Rather, the current version of NAPLAN, according to its developers, was built to reflect common content and skills addressed by the states and territories tests that preceded NAPLAN. This limitation is being addressed with a review of the National Assessment Program (NAP) subsequent to the release of the Australian Curriculum. The review aims to ensure the NAP, including NAPLAN, is aligned with the curriculum and provides the objective information necessary to drive continued improvements in student outcomes.

As its framers and developers admit, NAPLAN is a consensus test that focuses on basic skills. Stakeholders in some jurisdictions feel that it is not as rigorous as their state or territory tests that preceded NAPLAN. For example, all of NAPLAN's scales, except for writing, are composed of multiple choice and short answer items, replacing tests in some jurisdictions, according to stakeholders, that included extended response items addressing more complex thinking and deeper understanding.

There are at least two important alignment issues here. First, the logic of accountability and the value of data for improving instruction rest on the alignment between standards, curriculum and instruction, and assessment. If the assessments do not well match the learning goals then results have little value in judging how well students are learning, nor do the data have optimal value for diagnosing school or student needs if the assessments do not well match what students are expected to learn.

A second alignment issue is that NAPLAN's current focus on literacy and numeracy skills and lack of attention to so-called 21st century skills also limits its value in promoting Australia's education goals. The Melbourne Declaration (MCEETYA, 2008) clearly lays out an ambitious set of goals for all young Australians: "That all young Australians become successful learners, confident and creative individuals, and active and informed citizens." NAPLAN addresses a relatively narrow range of learning goals relative to what parent, teacher, principal, and business representatives with whom we spoke want for students. Clearly, this is not a problem unique to NAPLAN – there is a limit to what any time-limited, standardised test can address – but it is a potential concern if the system were to overemphasise NAPLAN results. While other components of the National Assessment Program may address other learning goals, for example periodic tests in ICT literacy, science literacy, civics and citizenship, the frequency and visibility of NAPLAN makes it a more important driver for Australia's educational system.

There are some challenges in ensuring NAPLAN is a fair test for some subgroups

The Melbourne Declaration (MCEETYA, 2008) makes clear the Australian Government's goals to promote equity and excellence. It gives schools responsibility for building on local cultural knowledge and experience and to reduce the effect of educational disadvantage, including socio-economic status, Indigenous status and disability. In the process of developing NAPLAN, it is unclear whether accessibility for Indigenous students and students with disabilities has received sufficient attention. The tension between standardised tests that are supposed to be common across all students and the need to be sensitive to local cultural knowledge and experience is a difficult one. However, technical analyses of the 2008 NAPLAN results (ACER, not dated) suggest that the tension has not yet been adequately balanced for Indigenous students. That is, differential item functioning analyses suggest that items are functioning differently for Indigenous students than for non-Indigenous ones and available data suggest a floor effect, *i.e.* achievement scales do not fully capture the lowest ends of the range and cannot well differentiate students at these levels. This is in spite of considerable efforts to design tests inclusive of all students. In particular, the NAPLAN test development process involves the consultation of Indigenous experts in states and territories. They provide specific feedback on the suitability of the test items and the appropriateness of the stimulus materials for Indigenous students. Test trials are also carried out using a sample of students, and analysis of the results is undertaken to ensure that all items are culturally appropriate and free of bias. However, while Indigenous representatives with whom we spoke were positive about the Australian Government's commitment to reducing the achievement gap, they still expressed concern for the fairness of the test and were not fully aware of NAPLAN efforts to eliminate cultural bias.

While the Melbourne Declaration speaks to equity for students with disabilities, it is not clear how far this principle has been carried for students with disabilities in the development of NAPLAN. Test accommodations exist in line with what students are normally allowed in the classroom but the consistency with which they are administered is unclear. In addition, attention to test functioning for students with disabilities was not included in the 2008 draft technical report (ACER, not dated). It should be noted, however, that such analyses were conducted for students with a language background other than English (LBOTE) and, with the exception of writing, the literacy and numeracy assessments were found psychometrically suitable for these students.

The challenge of timeliness in the delivery of NAPLAN results

While teachers and principals generally found NAPLAN a credible measure of basic skills, they observed that the timeliness of results limited its utility – *i.e.* the test is administered in the autumn, and results not provided until the following spring. They wanted the results back faster to inform their planning. The feasibility of their request, however, is moot, given tests for all students in the country at sampled grades need to be processed simultaneously and the desire to include more extended response options in NAPLAN to capture deeper learning and problem solving.

Some challenges in A-E reporting

The credibility of A-E reporting was a concern for some parent representatives, who were having difficulty adjusting their states' defining of a score of C (or equivalent), generally meaning "meets standard." These representatives saw "C" as a more negative

rating. Some, including parent and school representatives, also resisted the application of grades to the early years of schooling. They were concerned about the effects of grades on young children's self-concepts and the possible negative consequences of labelling students.

The primary purpose of a summary of the A-E reporting type is to inform students and their parents about performance against standards, learning potential and return to study effort. Research shows that this type of summary feedback to students may have a positive or negative influence on students' motivation and performance. Potentially, the information can raise student motivation and communicate to the student that his or her work is recognised and worth the effort. However, grades can also negatively affect students' motivation and effort, if the information conveyed communicates to the student low returns to effort or hurts his or her self-confidence (Sjörgen, 2009). In their review of the literature on formative assessment, Black and Wiliam (1998) found that the grading function in schools tends to be overemphasised while learning is underemphasised. Grades tend to shift attention away from feedback and improvement from mastery learning to extrinsic performance goals.

A major challenge is to align A-E ratings to the Australian Curriculum, an undertaking which has now started under the leadership of ACARA. This will bring a national agreement on A-E definitions improving the current situation where A-E definitions differ across states and territories. Another challenge is to ensure that teachers develop capacity to assess against A-E ratings (and against the Australian Curriculum, once the alignment with A-E ratings has been achieved). This will require teachers to find the Australian Curriculum concrete enough to guide their instruction and assessment, to be able to consistently interpret learning goals and to benefit from clear grading criteria.

Some inadequacies in teachers' skills for assessment and to use assessment data

In our meetings with stakeholders, there was indication of some inadequacies in teachers' preparation for student assessment. For example, teachers noted the limitations of some teachers' assessment knowledge and skills as they entered the profession. New teachers needed substantial support to acclimate to A-E reporting schemes and moderation processes, which apparently were not priority areas during teacher education programmes. A survey of Australian teachers reveals that "methods for assessing student learning and development" were among the areas of greatest need for professional development as identified by teachers: in primary schools, 65% of teachers indicated either a major or a moderate need (the 2nd area of greatest need among 16 identified areas, below "making more effective use of computers in student learning"); while in secondary schools, 55% of teachers expressed the same need (the 3rd area of greatest need among 16 identified areas) (McKenzie *et al*., 2008).

Also, the utility and sound use of data, of course, depends on teachers' assessment literacy and ability to appropriately integrate assessment data and learning in classroom instruction, including the appropriate use of standardised tests such as NAPLAN. During the Review visit, it was noted that many teachers, including beginning teachers, needed considerable support to analyse and interpret student assessment data and to reflect it into adjustments to classroom instruction. Representatives of teacher education institutions with whom we spoke mentioned standardised testing such as NAPLAN being a relatively new phenomenon in Australia and noted uneven attention in their programmes to teachers' capacity to understand, analyse and use standardised test data. This is in spite of training provided in most jurisdictions to improve the competency of teachers to analyse

and interpret student assessment data. For example, the Victorian Curriculum and Assessment Authority conducts in-service courses in schools around Victoria, to develop school leaders' and teachers' skills in interpreting results of NAPLAN and the Victorian Certificate of Education exam.

Consistency of secondary certificates across jurisdictions is not ensured

There are concerns about the consistency of secondary certificates across jurisdictions as no apparent processes of moderation are in place for cross jurisdiction comparison. Assessment and certification procedures at the end of Year 12 are specific to states and territories and these are typically moderated within and across schools but just for schools within a single jurisdiction.

There are risks the emphasis on NAPLAN may "narrow" teacher-based assessment

There are indications that NAPLAN is becoming dominant in discussions around "student assessment". While the Australian Government's student assessment framework clearly sees teacher-based summative and formative assessment as important elements, stakeholders often see these as less significant components of the explicit assessment system and do not seem to make the link between A-E reporting, formative assessment and NAPLAN results. When asked about "student assessment", stakeholders were most likely to talk about NAPLAN. Moreover, teachers and their representatives expressed concerns that the national focus on NAPLAN was drowning out attention to classroom-based summative and formative assessment, which they see as more critical for improving student learning. Admittedly the timing and scope of the Review may have influenced the perceptions that stakeholders shared. The Review Team visited shortly after a highly publicised proposed ban on NAPLAN and the national focus of the Review may have encouraged stakeholders to emphasise national assessment initiatives rather than longstanding state and territory practices. Nonetheless, that NAPLAN is given annually and is linked to federal funding and school reporting makes it a highly visible assessment that is likely to send a strong signal to administrators, teachers and students about what is most important for teaching and learning.

Research shows that while summative assessment is primarily conceived to measure the outcomes of learning, the approach to summative assessment can in turn have a strong impact on the learning process itself. Different assessment policies and practices influence students' motivation, effort, learning styles and perceptions of self-efficacy as well as teaching practices and teacher-student relationships (Nusche, forthcoming). In particular, the impact of summative assessment on the scope and depth of teaching and learning activities can be positive if the assessment signals and clarifies the full range of goals that students are expected to achieve. But if the scope of summative assessment only covers a small fraction of the overall curriculum goals, then the impact of assessment on teaching and learning can be restrictive (Harlen, 2007). This underlines the importance of student assessment which encompasses the entirety of student learning objectives such as that conducted by teachers in their classrooms.

Policy recommendations

Develop national consistency while respecting state and territory assessment strengths and cultures

All stakeholders with whom the Review Team spoke understood and supported the need for a national curriculum and supported the move to national standards and assessments. This support is a tribute to the collaborative process in which the Australian Government engaged and to the level at which NAPLAN is addressed: that is, basic skills and predominantly pre-secondary education, where there is not a strong or long tradition of assessment in states and territories. However, the Review Team witnessed considerable variation in prior assessment history, capacity and culture across the states and territories. Those with the strongest history in pre-secondary testing were concerned that NAPLAN and its reporting were less comprehensive and sophisticated than the state or territory tests that had preceded it. Policy makers and ACARA should consider how the MCEECDYA's annual assessment program can take advantage of and nourish the unique strengths and capacities available in various states and territories, for example, by linking or otherwise integrating the *My School* website data with locally available data and by sharing and refining advanced assessment resources, strategies and tools that have been locally developed.

At the secondary level, the Review Team was impressed by the rigour and culture supporting the assessment system. While systems varied from jurisdiction to jurisdiction, each was well established and teachers and students appeared highly invested and engaged in their current efforts. Stakeholders in some states and territories were concerned that the systems, relationships and local cultures that are currently working very well would be sorely disrupted and educational programmes and values compromised if the Australian Government attempts to put a standardised, secondary assessment system into place. Policy makers should consider how the strengths of diverse current secondary assessment systems can be accommodated rather than homogenised. For example, policy makers may want to consider how moderation of secondary assessments across jurisdictions can best support comparability of senior secondary assessments rather than enforcing a single system, an analysis which fits well ACACA's responsibilities. This concern is also likely to be lessened as the Australian Curriculum is introduced in the senior secondary years (Years 11 and 12). Comparability of secondary assessments/certificates within and across jurisdictions will then be supported through the development of teachers' capacity to grade against the curriculum, grading criteria and exemplars illustrating student performance, in combination with moderation processes (see below).

Reinforce the assessment validity of NAPLAN

While the Review Team found strong evidence of NAPLAN's reliability and lack of bias for classifying non-Indigenous and non-disabled students relative to its five national achievement scales (reading, writing, spelling, grammar and punctuation, and numeracy), it found less attention to some important validity issues. First and foremost is NAPLAN's alignment with the Australian Curriculum and the extent to which NAPLAN is balanced in its representation of the depth and breadth of intended student learning goals. Because NAPLAN preceded the introduction of the Australian Curriculum, this alignment challenge was inevitable, and ACARA and NAPLAN's developers, as we have noted, surely are aware of the problem and have plans to alleviate it.[3] However, in doing so

assessment authorities will face important tensions: NAPLAN currently is framed as a basic skills test; but more fully representing the intended goals of the Australian Curriculum may require a broader orientation, especially given the high visibility of NAPLAN results and the potential danger of curriculum narrowing. In any event, to serve as a worthy target for teaching and learning and an accurate measure of students' attainment of curriculum goals, NAPLAN's alignment with those goals needs to be assured. This effort could be supported by independent studies of alignment.

Evidence of instructional sensitivity and transfer are additional areas that may be of interest. Holding schools accountable and evaluating school agents based on their students' test results requires evidence that student test scores indeed are influenced by good teaching and sound instructional interventions and are not simply a function of test preparation and/or innate ability.

The specific purposes that NAPLAN may be expected to serve each bring implicit requirements for additional validity data, and requirements for serving some purposes may be at odds with others. For example, measures of student growth require tests that address a consistent set of targets over time and are vertically scaled, as NAPLAN is, which tends to narrow the breadth of content that can be assessed. Vertically scaled tests and comprehensive measures of the breadth and depth of student learning, in short, tend to be opposing goals. Australia's policy makers need to recognise the tensions in the various uses of NAPLAN and keep clear priorities so that the test can be appropriately designed and valid evidence collected.

Ensure that NAPLAN is a fair test to all subgroups

NAPLAN's accessibility and lack of bias for Indigenous students and students with disabilities also merits attention. Available evidence suggests challenges in how the test functions for Indigenous students, and data on test functioning for students with disabilities is largely absent. Test developers need to consider how test design and development may mitigate such shortcomings and continue to gather evidence to document their success in this area.

In ensuring that NAPLAN is a fair test for Indigenous students, continuing to involve Indigenous representatives in the initial test specification and item and test development process will help increase the cultural sensitivity of the test and help alleviate perceptions of potential bias. This should be combined with a good communication of NAPLAN's efforts to eliminate cultural bias so results are trusted among the Indigenous population.

In the same way, test accommodations for students with disabilities need to be reinforced to more systematically reflect the kind of support and assistance the students usually receive in the classroom. While accommodations are available, the consistency with which they are administered is an open question, and validity data are largely missing. The objective is to make NAPLAN an inclusive test based on the principle that all students have the opportunity to participate in educational activities, including assessment activities, and to demonstrate their knowledge, skills and competencies.

Establish safeguards against overemphasis on NAPLAN

Parents, teachers, principals and administrators expressed concern that the high visibility of NAPLAN results was encouraging schools and educators to narrow the curriculum to the basic skills addressed by the current test, at the expense of knowledge

and skills that are not annually assessed or reported at the individual student and school levels, including science and history as well as the complex thinking, problem-solving, collaboration and networking, and ICT skills they feel are essential to the children's future success. Expanding the range of indicators reported and used to judge schools will help to mitigate this narrowing, but policy makers may want to consider what other safeguards and cross-checks they can put in place to reduce this threat (see also Chapter 5). It should be noted that the Review Team heard a number of anecdotes about potential negative effects of NAPLAN, but was unable to find documented research evidence. Ongoing evaluation of assessment use and consequences can help programme designers to maximise its productive use and minimise undesirable consequences.

Strengthen teachers' capacity to assess student performance against the Australian Curriculum

In Australia's standards-based system, and in particular following the introduction of the Australian Curriculum, sound strategies to assess against the standards/curriculum are paramount. The current strategy for student assessment consists of a combination of NAPLAN, limited to measure a subset of student learning goals, and teacher-based assessments against the full range of curriculum goals (and reflected in A-E reporting). The latter implies a considerable investment on teacher capacity to assess against the standards, including specific training for teachers, the development of grading criteria and the strengthening of moderation processes within and across schools. This will be facilitated by the alignment of A-E ratings to the Australian Curriculum, an area of priority which is currently receiving attention through work led by ACARA. This work will bring the desirable consistency of A-E definitions across states and territories and will assure the proper link between teacher-based student assessment and the Australian Curriculum.

It is essential that the development of the Australian Curriculum (and its alignment with A-E ratings) is followed by a considerable investment on teachers' capacity to assess against its objectives to ensure the reliability of teachers' A-E reporting. This could be a priority for professional development activities of teachers in the coming years. Training should include a range of aspects such as the ones illustrated in Box 3.1 – for instance, the ability to understand different aspects of validity – what different assessments can and cannot reveal about student learning – or strategies to ensure the inclusiveness and fairness of an assessment. Educational authorities may also want to strengthen the development of additional tools to support teacher assessment, such as exemplars illustrating student performance at different levels of achievement, and scoring rubrics listing criteria for rating different aspects of performance. This can help guide A-E reporting. Box 3.2 describes several strategies to improve the reliability of teacher-based summative assessment.

Australia's policy makers may want to consider how they can assure that assessment capacity is reflected in teacher standards and addressed in teacher preparation programmes and how they can encourage teacher collaboration and states and territories sharing of capacity building resources in this area.

Box 3.2 Strategies to improve the reliability of teacher-based summative assessment

The research literature describes several ways to address potential bias in teachers' summative assessments and to help make these assessments more reliable.

Scoring guides. Teacher-based assessment can be facilitated by providing teachers with scoring guides including detailed descriptions of competency levels and examples of high performance. Such scoring criteria should be detailed but generic so that they can be adapted to the full range of classroom work (Harlen, 2004). Two studies from the United States showed that teacher-based assessment of science projects was highly reliable when teachers used detailed scoring criteria (Frederiksen and White, 2004; Shavelson *et al.*, 1992, in Harlen, 2004). In a study from Victoria, teachers used "subject profiles" to rate student achievement in relation to indicators of different bands of achievement for each component of each subject in the curriculum. The study indicates that teachers made reliable judgments using these indicators (Rowe and Hill, 1996, in Harlen, 2004).

Teacher participation in the development of criteria. Teacher involvement in developing criteria to score student achievement can strengthen the reliability of their assessment. A strand of research points to the fact that teachers apply assessment criteria more consistently and accurately if they are clear about the goals to be achieved, and especially if they have participated in the development of criteria to score student achievement (Hargreaves *et al.*, 1996; Frederiksen and White, 2004).

An external yardstick. It is also essential that teachers have external benchmarks showing what is expected to be "normal" or "adequate" performance of students in a particular grade and subject. In Sweden, for example, teachers are encouraged to compare the achievements of students in internal assessment to student results in national tests and make corrections where there are major discrepancies. In this context, the entire responsibility for student grading rests with the teachers themselves, but via the national tests, teachers are given a tool to compare their own assessment to an external guidance and reference point. When determining the final grade, teachers are encouraged to take all available information – including the national test results – into consideration.

Training for teachers. Training for teachers might help avoid error and bias in assessment. Training should include how to identify valid evidence, as well as how to apply grading criteria to very different types of evidence of student learning. Training can also focus on making teachers aware of their own potential unconscious bias in making judgments about different groups of students. Training can strengthen teachers' assessment literacy, which includes awareness of the factors that influence validity and reliability of results, and capacity to make sense of data and track progress (Earl and Fullan, 2003; Fullan, 2001).

Teacher collaboration in assessment. Some authors argue that reliability can also be developed without the use of externally designed standardised tests, focusing instead more on using multiple human judgments (Van der Vleuten and Schuwirth, 2005, in Baartman *et al.*, 2006). In Sweden, in many schools it is encouraged that teachers work together to grade each others' students rather than relying only on their own judgment (Nusche *et al.*, 2011).

External moderation. Several authors point out that external moderation can help correct errors and bias in teacher-based assessment. According to Somerset (1996), teacher-based assessment can only play an important role in quality evaluation if mechanisms are available to measure differences between teachers and control their effects.

Source: Adapted from Nusche (forthcoming).

Build teachers' competence to use student assessment data

A priority is to develop teachers' capacity to use student assessment data, including that generated by NAPLAN, for the improvement of classroom instruction. Teachers often report several positive effects of using student test results: greater differentiation of instruction, greater collaboration among colleagues, increased sense of teacher efficacy and improved identification of students' learning needs (Barneveld, 2008). This calls for the provision of formal training, possibly as a professional development option for teachers, on skills for analysing and interpreting student assessment data. Similarly, initial teacher education institutions should be encouraged to give sufficient attention to the formative use of NAPLAN results in their programmes.

Teachers' analysis and use of NAPLAN results to diagnose student learning needs represents another case in point. Sound diagnosis requires reliable diagnostic data. Yet, NAPLAN developers make clear that results are reliable only at the level of the five national achievement scales: reading, writing, spelling, grammar and punctuation, and numeracy. In fact, developers conducted studies to determine whether it was possible to provide reliable subscores to support individual diagnosis in numeracy and reading. They concluded it was not possible: the scores would not be sufficiently reliable for individual decision making. At the same time, tools enable teachers to use student responses to an individual item to diagnose student needs. In the absence of other sources of information, such inferences are not reliable and may well yield inaccurate diagnoses. This is reflected in NAPLAN documentation which states that it should be used to supplement other assessment information gathered by the teacher. The general issue is how ACARA can best communicate both what uses are justified, and which are not and should be discouraged. Some jurisdictions seem to have such strategies in place, while others do not. Another issue is to ensure that NAPLAN results are instructionally useful, including the diagnostic value of reports and the timeliness in the delivery of the results – advances in automated scoring may help in this latter respect.

Maintain the centrality of teacher-based assessment while ensuring the diversity of assessment formats

The current prominence of NAPLAN within the student assessment framework requires particular care about not reducing the importance of teacher-based assessment. As explained by Nusche (forthcoming), several studies underline that teacher-based summative assessment has a greater potential to improve approaches to teaching and learning than external tests. Teacher-based assessment takes place throughout the course and generally the work is returned to students along with feedback on strengths and weaknesses (Crooks, 2004). This type of assessment thus provides opportunities for teachers to adapt instruction and for students to adjust their learning styles and improve results. Also, since teachers are able to assess students' progress toward the full range of goals set out in the curriculum over time and in a variety of contexts, their assessments help to increase the validity of assessment (Harlen, 2007). Stronger assessment roles for teachers may also help to build their assessment literacy and skills and strengthen their professionalism (Looney, 2011).

However, it needs to be recognised that teacher-based assessments are often perceived as unreliable. Test items and grading standards may vary widely between teachers and schools, so that the results of teacher-based assessment will lack external confidence and cannot be compared across schools. There might also be a high risk of bias in teacher-based assessment, *i.e.* the assessment is unfair to particular groups of

students (even if there are strategies to address the reliability of teacher-based assessment, as described in Box 3.2). This indicates that there is a case for combining teacher-based assessment with external assessment, which tends to be more reliable (see Nusche, forthcoming), especially when stakes for students are high. Another approach is to develop on-demand assessments, where teachers can draw from a central bank of assessment tasks and ask students to take the assessment when they consider that they are ready.

Research indeed suggests that providing multiple opportunities and formats for student assessment increases both the validity and reliability of authentic assessment (Linn *et al.*, 1991). As explained by Nusche (forthcoming), complex assessments that combine different formats can balance the reliability of metric-based or standardised assessments with the validity of performance assessments (which assess a range of integrated knowledge and skills by asking students to perform a task). The right balance of assessment formats will partly depend on the purpose of assessment. The higher the stakes of an assessment, the more important it is that the assessment is highly reliable. A summative assessment with low stakes, such as sharing a judgement on student performance within school or with parents will place more importance on high validity, whereas a summative assessment used to determine access of students for higher education will focus more on reliability (Blok *et al.*, 2002; Harlen, 2007).

Offering a range of different assessment formats and tasks is also important for ensuring fairness in assessment. Several studies report that certain formats of assessment may advantage or disadvantage certain groups of students (Nusche, forthcoming). In England it is reported that girls tend to achieve higher scores than boys on open-ended tasks, whereas this gap narrows when multiple choice tests are used (Gipps and Murphy, 1994, in Gipps and Stobart, 2004). Assessments that place great emphasis on written tasks may disadvantage students from cultural traditions in which oral communication is prevalent (Rudduck, 1999, in Gipps and Stobart, 2004).

Increase the visibility of the Australian Government's goals for formative assessment

While research shows that formative assessment is a powerful lever for improving student learning, the Review Team found relatively little explicit attention to it in policy documents or discussion. Policy makers may want to consider how they make their goals for the use of formative assessment more explicit and take action to assure teachers' capacity to effectively engage with it. In some states and territories, such tools and strategies are in place and could be shared, but policy makers need to be aware and invest in the re-engineering necessary to align current formative assessment practices with the Australian Curriculum goals.

Effective formative assessment requires that teachers develop sophisticated skills for uncovering students' level of understanding, for providing feedback and adjusting teaching strategies to meet identified needs, and for helping students to develop their own skills for learning to learn. Strategies to improve the impact of formative assessment might include a stronger focus on short-cycle classroom interactions, building teachers' repertoire of research-based formative assessment techniques, and strengthening the approaches to respond to identified learning needs (OECD, 2005).

A student-centred approach to learning and assessment is also critical in ensuring the effectiveness of formative assessment. This includes engaging students in setting goals for learning, and developing students' skills for self- and peer-assessment. Box 3.3 illustrates Sweden's focus on engaging students in learning through Individual Development Plans (IDP). The IDP, as a core feature of Swedish education, ensures that both teachers and students are focused on identifying individual learning goals, and developing strategies to address any shortcomings (Nusche *et al.*, 2011).

Box 3.3 Individual Development Plans for students in Sweden

In Sweden, formative assessment is supported by Individual Development Plans (IDP) for students. Individual school leaders set out the general template for the IDP that will be used in their school. The IDP is to include an assessment of the student's current performance levels in relation to learning goals set in the curriculum, and steps the student should take to reach those goals. Whether to include additional information, such as the student's more general development (*e.g.* the student's ability to take on responsibility, their social skills) is up to the school leader. The written IDP is to include the student's and guardian's input from the regular development talks, which usually take place once a semester. For students who are experiencing difficulty, schools are required to document plans as to how they will help students achieve goals.

The IDP, as a core feature of Swedish education, ensures that both teachers and students are focused on identifying individual learning goals, and developing strategies to address any shortcomings. It can be a powerful tool for developing students' own assessment skills, as well.

Source: Nusche *et al.* (2011).

Notes

1. The "Frequency of student assessment by method" indicator in Annex D provides an overview of student assessment methods used for 15-year-olds according to the 2009 PISA survey.
2. According to the 2009 PISA survey, 81.7% of 15-year-old students were in schools where the principal reported that teachers "have considerable responsibility in establishing student assessment policies" (8th highest figure among OECD countries against an average of 69.0%) (see Annex D).
3. ACARA has been mandated to review the National Assessment Program and its alignment with the curriculum once the Australian Curriculum has been fully implemented.

References

ACER (Australian Council for Educational Research) (n.d.), *NAPLAN 2008 Central Analysis Technical Report*, Camberwell, Victoria.

AERA, APA, NCME (American Educational Research Association, American Psychological Association, National Council for Measurement in Education) (1999), *Standards for Educational and Psychological Testing*, AERA, Washington D.C.

Australian Government (2010), *Country Background Report for Australia*, prepared for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, Department of Education, Employment and Workplace Relations, Canberra, available from *www.oecd.org/edu/evaluationpolicy*.

Baartman, L.K.J., T.J. Bastiaens, P.A. Kirschner and C.P.M. Van der Vleuten (2006), "The Wheel of Competency Assessment: Presenting Quality Criteria for Competency Assessment Programs", *Studies in Educational Evaluation*, Vol. 32, pp. 153-170.

Baker, E., R. Linn, J. Herman and D. Koretz (2002), "Standards for Educational Accountability Systems", *CRESST Policy Brief No. 5*, National Center for Research on Evaluation, Standards and Student Testing, Los Angeles, California.

Barneveld, C. (2008), "Using Data to Improve Student Achievement", *Research Monograph No. 15*, The Literacy and Numeracy Secretariat, Ontario, *www.edu.gov.on.ca/eng/literacynumeracy/inspire/research/Using_Data.pdf*.

Black, P. and D. Wiliam (1998), "Assessment and Classroom Learning", *Assessment in Education*, 4, 7-74.

Blok, H., M.E. Otter and J. Roeleveld (2002), "Coping with Conflicting Demands: Student Assessment in Dutch Primary Schools", *Studies in Educational Evaluation*, Vol. 28, pp. 177-188.

Crooks, T.J. (2004), "Tensions between Assessment for Learning and Assessment for Qualifications", paper presented at the Third Conference of the Association of Commonwealth Examinations and Accreditation Bodies (ACEAB), Fiji, March 2004.

Earl, L. and M. Fullan (2003), "Using Data in Leadership for Learning", *Cambridge Journal of Education*, Vol. 33, pp. 383-394.

EPPI (2002), *A Systematic Review of the Impact of Summative Assessment and Tests on Students' Motivation for Learning*, Review conducted by the Assessment and Learning Synthesis Group, Evidence for Policy and Practice Information and Co-ordinating Centre, London.

Frederiksen, J. and B. White (2004), "Designing Assessment for Instruction and Accountability: An Application of Validity Theory to Assessing Scientific Inquiry", in (ed.) M. Wilson, *Towards Coherence between Classroom Assessment and Accountability*, 103rd Yearbook of the National Society for the Study of Education, part 11, National Society for the Study of Education, Chicago.

Fullan, M. (2001), *Leading in a Culture of Change*, Jossey-Bass, San Francisco.

Gipps, C. and G. Stobart (2004), "Fairness in Assessment", in GTC, *Perspectives on Pupil Assessment*, General Teaching Council (GTC) for England, London.

Gipps, C. and P.M. Murphy (1994), *A Fair Test?*, Open University Press, Buckingham.

Hargreaves, D.J., M.J. Galton and S. Robinson (1996), "Teachers' Assessments of Primary Children's Classroom Work in the Creative Arts", *Educational Research*, 38, pp. 199-211.

Harlen, W. (2004), "Can Assessment by Teachers be a Dependable Option for Summative Purposes?" in GTC, *Perspectives on Pupil Assessment*, General Teaching Council (GTC) for England, London.

Harlen, W. (2007), "Criteria for Evaluating Systems of Student Assessment", *Studies in Educational Evaluation*, Vol. 33, 15-28.

Herman, J. (2010), "Toward Coherent Assessment Systems", *CRESST Policy Brief*, National Center for Research on Evaluation, Standards and Student Testing, Los Angeles, California.

Linn, R.L., E.L. Baker and S.B. Dunbar (1991), "Complex, Performance-based Assessment: Expectations and Validation Criteria", *Educational Researcher*, 20.

Looney, J. (2011), "Integrating Formative and Summative Assessment: Progress Toward a Seamless System?", paper commissioned for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, *OECD Education Working Papers*, No. 58, OECD, Paris, available from *www.oecd.org/edu/workingpapers*.

MCEECDYA (2009), *Principles and Protocols for Reporting on Schooling in Australia*, Ministerial Council for Education, Early Childhood Development and Youth Affairs, *www.mceecdya.edu.au/mceecdya/nap_principles__protocols_for_rep_on_school_2009,27896.html*.

MCEETYA (2008), *Melbourne Declaration on Educational Goals for Young Australians*, Ministerial Council on Education, Employment, Training and Youth Affairs, Canberra.

McKenzie, P., J. Kos, M. Walker, J. Hong and S. Owen (2008), "Staff in Australia's Schools 2007", *Teaching and Learning and Leadership, http://research.acer.edu.au/tll_misc/3*.

McLaughlin, M. and J. Talbert (2006), *Building School-based Teacher Learning Communities: Professional Strategies to Improve Student Achievement*, Teachers' College Press, New York.

Nusche, D. (forthcoming), "Summative Assessment: What's in it for Students?", *OECD Education Working Paper Series*, OECD, Paris, to be available from *www.oecd.org/edu/workingpapers*.

Nusche, D., G. Halász, J. Looney, P. Santiago and C. Shewbridge (2011), *OECD Reviews of Evaluation and Assessment in Education: Sweden*, OECD, Paris, available from *www.oecd.org/edu/evaluationpolicy*.

OECD (2005) *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD, Paris, *www.oecd.org/document/39/0,3343,en_2649_35845581_34365415_1_1_1_1,00.html*.

Rowe, K.J. and P.W. Hill (1996), "Assessing, Recording and Reporting Students' Educational Progress: The Case for 'Subject Profiles'", *Assessment in Education*, 3, 309-352.

Rudduck, J. (1999), Seminar presentation, QCA, London, as cited in Gipps and Stobart (2004).

Sadler, D.R. (1989), "Formative Assessment and the Design of Instructional Systems", *Instructional Science*, 18(2), 119-144.

Sebba, J. and G. Maxwell (2005), "Queensland, Australia: An Outcomes-based Curriculum", in *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD, Paris, pp. 191-204.

Shavelson, R.J., G.P. Baxter and J. Pine (1992), "Performance Assessments: Political Rhetoric and Measurement Reality", *Educational Researcher*, 21, 22-27.

Sjörgen, A. (2009), "The Long Run Consequences of being Graded in Elementary School", paper presented at the Final Conference of the RTN Network "Economic of Education and Education Policy in Europe" (EEEPE), hosted by the Centre for Economic Performance, LSE, London.

Somerset, A. (1996), "Examinations and Educational Quality", in A. Little and A. Wolf, *Assessment in Transition, Learning Monitoring and Selection in International Perspective*, Pergamon, Elsevier Science Ltd., Oxford.

Vleuten, C.P.M. van der, and L.W.T. Schuwirth, (2005), "Assessing Professional Competence: From Methods to Programmes", *Medical Education*, 39, 309-317.

Wiliam, D. and M. Thompson (2007), "Integrating Assessment with Instruction: What will it Take to Make it Work?", in C. Dwyer (ed.), *The Future of Assessment: Shaping Teaching and Learning*, Lawrence Erlbaum Associates, Mahwah, New Jersey.

Chapter 4

Teacher appraisal

Teacher appraisal varies across states and territories but typically occurs in three specific instances: (i) To gain registration to teach within the state or territory; (ii) As part of the employer's performance management processes; and (iii) To gain promotion positions in schools in recognition of quality teaching performance (Advanced Skills Teaching positions). Particularly positive features of teacher appraisal include the existence of teaching standards; registration processes which are consolidated; performance management processes which provide a good basis for developmental teacher appraisal; and Advanced Skills Teaching positions which grant opportunities for recognition of skills and competencies. Priorities for future policy development include aligning teaching standards with a competency-based career structure for teachers; conceiving teacher registration as career-progression evaluation; performing developmental evaluation through teacher appraisal as part of performance management processes; reinforcing linkages between teacher appraisal, professional development and school development; and strengthening competencies for teacher appraisal. These policies seek to render teacher appraisal more systematic and meaningful across the system; provide teachers with more opportunities for feedback; better address cases of underperformance; better align competencies at different stages of the career and the roles and responsibilities of teachers in schools; and improve the recognition of teachers' work.

This chapter looks at approaches to teacher appraisal within the Australian overall evaluation and assessment framework. Teacher appraisal refers to the evaluation of individual teachers to make a judgement about their performance. Teacher appraisal has typically two major purposes. First, it seeks to improve teachers' own practices by identifying strengths and weaknesses for further professional development – the improvement function. Second, it is aimed at ensuring that teachers perform at their best to enhance student learning – the accountability function (Santiago and Benavides, 2009). An overview of the main features of the teaching profession in Australia is provided in Box 4.1.

Context and features

Teacher appraisal procedures

Teacher appraisal procedures vary across states and territories but typically occur in three specific instances:

- To gain registration/accreditation to teach within the state or territory;
- As part of the employer's performance management processes; and
- To gain promotion positions in schools in recognition of quality teaching performance (Advanced Skills Teaching positions).

This is in addition to probationary period processes which are common in Australian schools. Most employers require a summative evaluation at the end of the 1- or 2-year probationary period. In general the principal takes responsibility for this evaluation which involves teachers providing evidence of effective teaching such as lesson plans and portfolios.

Teacher registration/accreditation

Registration is a requirement for teachers to teach in Australian schools, regardless of school sector. All states and territories, with the exception of the Australian Capital Territory (ACT),[1] have existing statutory teacher registration authorities responsible for registering teachers as competent for practice. The levels of teaching accreditation vary according to the jurisdiction (see Table 4.1). In most jurisdictions, teachers reach the first level of accreditation from the relevant authority upon graduation from an approved initial teacher education programme. Each of the teacher registration authorities has its distinct set of standards for registration/accreditation, but these are overall comparable (see Table 16 in Australian Government, 2010).

Table 4.1 Levels of teacher registration/accreditation across Australia

Registration authority	Levels of registration/accreditation			
New South Wales Institute of Teachers	Graduate Teacher	Professional Competence	Professional Accomplishment	Professional Leadership
Teacher Registration Board of the Northern Territory	Graduate teacher	Competent Teacher	Accomplished Teacher	
Queensland College of Teachers	Provisional Registration	Full Registration		
Teachers Registration Board of South Australia	Graduate	Full Registration		
Teachers Registration Board of Tasmania	Graduate	Competence	Accomplishment	Leadership
Victorian Institute of Teaching	Graduating Teacher	Full Registration		
Western Australian College of Teaching	Provisional Registration	Full Registration		

Source: Reproduced from Australian Government (2010).

Advancement to full registration (or professional competence) is achieved after a period of employed teaching practice – which varies across jurisdictions from 80 days in Victoria to 200 days in Queensland, South Australia and Western Australia – and an appraisal against full registration standards. Approaches to the latter vary across states and territories (see Table 13 in Australian Government, 2010). In South Australia, the principal rates the applicant on a five-point scale (unsatisfactory to outstanding) against elements of the teaching standards and makes a recommendation. In Queensland the principal prepares a "Provisional to Full Registration Recommendation Report" based on examples of practice to demonstrate achievement of the standards. In the Northern Territory, Tasmania and Victoria, the applicant provides evidence of meeting the standards to a school-based panel while in Western Australia the applicant collects evidence demonstrating ability to meet the standards and is followed by a mentor who makes the recommendation for registration. In New South Wales, the Institute of Teachers makes the accreditation decision on the basis of a school-based evaluation of teaching practice against the standards. In New South Wales, Tasmania and the Northern Territory, teachers can achieve, on a voluntary basis, higher levels of registration (*Accomplishment* and *Leadership*) using the distinct set of standards describing each stage.

In all states and the Northern Territory, after teachers have initially become registered within their jurisdiction, they must renew their registration. The period of registration is commonly five years, with the exception of South Australia where it is three years, and Victoria where all teachers are required to renew their registration annually as of 1 January 2011. The process varies across jurisdictions but essentially consists of minimum requirements for participation in professional development activities (see Table 14 in Australian Government, 2010).

Performance management

Teacher appraisal conducted as part of regular employer's performance management processes varies considerably across jurisdictions and schools. Employers take responsibility for the implementation and management of their performance management processes and, in the case of government schools, these may be mandated under the terms of the jurisdiction's public service legislation. Table 4.2 provides an overview of the overall framework for teacher regular appraisal in government schools for each Australian jurisdiction. It is in general an annual process internal to the school but, in some cases, linked to external systemic processes such as overarching school improvement frameworks (for both the government and non-government sectors). The primary focus in all instances is as a supportive and development process to assist teachers in their professional career development. Records relating to each appraisal are generally not maintained centrally.

Teacher appraisal procedures are generally managed within the school by the principal or his/her nominee (typically the teacher's line manager or supervisor). The typical procedure, with considerable variation across schools in the extent to which each of the components of the process is applied, is the following. At the commencement of a new cycle, the teacher and his/her supervisor agree a performance plan; they meet mid-cycle to discuss progress against the plan; and then review performance against the plan at the end of the cycle. The performance plan intends to align the teacher's goals with his/her professional roles and responsibilities and with school and system priorities. Regular informal feedback is regarded as an integral part of the process. In the non-government sector, additional criteria may be identified by the school and teachers that specifically relate to their position or school's values and ethos.

The performance criteria and reference standards used in teacher appraisal draw mostly on teaching standards for the respective jurisdiction (even if, in some jurisdictions, these might be professional standards set up by the respective educational authorities instead of the registration standards developed by the jurisdiction's statutory teacher registration authority). Other reference documents typically include strategic educational plans for the respective jurisdiction and school plans.

The procedures employed in schools vary, but may include: classroom observation (mandatory in the Northern Territory), self-reflection, peer review, support of a mentor, formal coaching, appraisal meeting with assessment panel or supervisor, interview, teacher portfolio, student results or learning outcomes. According to the PISA 2009 survey, 41.3% of Australian 15-year-old students are in schools where the principal reported that student achievement data are used in the evaluation of teachers' performance (16th highest figure among OECD countries, against an average of 44.2%, see Annex D). In the Northern Territory, teachers may choose to use a Behaviourally Anchored Rating Scale (BARS) to rate their current performance. The BARS uses key behavioural descriptors of practice based on the Northern Territory Professional Standards for Teachers.

Table 4.2 Framework for teachers' performance management in government schools across Australian jurisdictions

State or territory	Framework for teachers' performance management
Australian Capital Territory	Performance management is undertaken around the teacher's Professional Pathways Plan, which is designed to support continuous improvement of performance and career goal setting. The plan is developed annually by a teacher in consultation with his or her professional mentor and identifies and records agreed goals and strategies for the school year. It is subject to formal review at least twice a year. The process has a focus on outcomes resulting from a teacher's performance, and features explicit and negotiated performance measures.
New South Wales	Teachers are subject to an annual performance review, usually led by the principal, as specified in the Industrial Relations Commission of New South Wales Crown Employees (Teachers in Schools and Related Employees) Salaries and Conditions Award 2009. Types of evidence expected to support the review are: conferences between the teacher and the principal; observations of educational programmes; and review of documentation such as lesson planning, lesson material and student work, and evaluations and reports. The performance review process is associated with the demonstration of "continuing efficiency in teacher practice", which is necessary for salary progress. The Teacher Assessment and Review Schedule includes the standards to assess and develop teacher performance in alignment with the NSW Institute of Teachers' Professional Teachers Standards.
Northern Territory	Teachers are expected to undergo an annual performance review process in accordance with requirements that apply to all NT public sector employees. Performance expectations are established in Performance Agreements and are based on actions identified in the Annual Operational Plan, as well as teaching standards. Performance data is a mandated part of the performance review. It is expected that a teacher's performance aligns with the needs and goals of the school under the Accountability and Performance Improvement Framework. Classroom observation is mandatory in the development of the teacher's performance plan. While teachers may choose the person who will observe them in the first instance, there is an expectation that the principal or another school leader will also observe the teacher.
Queensland	The Developing Performance Framework is being introduced in schools. The agreement between the Queensland Teachers Union and the Department of Education states that "in establishing and implementing the Framework, each school or work place should adopt an approach that is appropriate to its needs. Unlike traditional models of individual performance appraisal, the framework supports group, team, collegial and mentoring approaches to the process of developing performance". All employees (from teachers to principals) are expected to use the document as a framework to negotiate tasks and priorities for both the school and for the individual.
South Australia	Teachers are subject to SA's Partnerships for Performance policy, which applies to all state employees. It requires that all teachers develop an annual performance and professional development plan in co-operation with their line manager. A Quality Performance Development Pilot is being conducted during 2010 to work towards building capacity in the giving and receiving of feedback and the development of a performance improvement culture in sites.
Tasmania	Teachers are subject to an annual performance review, which is a necessary precondition for salary progression. As other public sector employees, teachers in government schools are required to create an annual Performance Plan based on their statement of duties.
Victoria	The Department of Education and Early Childhood Development mandates performance and development arrangements for all school staff within the context of the Victorian Government Schools Agreement 2008. The arrangements provide a framework for: *(i)* Review of performance; *(ii)* Feedback that supports ongoing learning and development; and *(iii)* A supportive environment for improving performance where the required standards are not met. The procedures for managing the performance of Victorian teachers are set out in the Victorian Performance and Development Guide. Performance and development includes an assessment on an annual basis for all teachers (with salary progression for eligible teachers) and is organised in four main stages: *(i)* Performance plans prepared and agreed with the principal; *(ii)* Mid-cycle review to discuss the teacher's progress; *(iii)* Assessment of the teacher's performance against the standards; and *(iv)* Performance plans prepared and agreed with the principal for the next cycle informed by the outcome of the last cycle. Appraisal for teachers is linked to the Department's own expected professional standards (distinct from those of the Victorian Institute of Teaching) – which describe the responsibilities of the three career stages – graduate, accomplished and expert teacher.
Western Australia	Teachers in government schools are subject to the same performance management requirements as all public sector employees. The Department provides no strict guidelines for individual performance management. The emphasis is on the role of the principal as taking responsibility for the assessment of performance.

Source: Australian Government (2010) and Jensen and Reichl (2011).

A range of strategies is used to implement performance management across the system, including:

- Training for line managers in schools;
- Central delivery of professional development programmes;
- Incorporation of performance management into departmental policies and training; and
- Provision of support material for schools – for example, templates, website resources.

In the non-government sector, performance management may not be mandated, and the number of schools with formalised programmes, the frequency of appraisal and the purpose of the process varies considerably.

Advanced Skills Teaching (AST) positions

Teachers can also undergo appraisal, on a voluntary basis, to gain promotion positions in schools in recognition of quality teaching performance (Advanced Skills Teaching positions). These positions carry higher pay and are generally associated with further responsibilities and specific roles in schools. In most cases, teachers do not have to be at the top of the salary scale to apply for these positions which entails a thorough assessment of their performance. Table 4.3 provides an overview of such schemes across Australia (see also Ingvarson *et al.*, 2007). Some similar schemes operate in the non-government sector (*e.g.* Experienced Teacher (Level 2) classification in Victorian Catholic Schools and Advanced Skills Teacher in Queensland Catholic schools).

Table 4.3 Advanced Skills Teaching schemes in government schools across Australian jurisdictions

State or territory	Promotion position	Description of scheme
New South Wales	Highly Accomplished Teacher (HAT)	NSW introduced the Highly Accomplished Teacher (HAT) position in July 2009. The HAT position is an initiative of the Smarter Schools National Partnership on Improving Teacher Quality. A HAT is an excellent teacher who models high-quality teaching for his/her colleagues across the school and leads other teachers in the development and refinement of their teaching practice to improve student learning outcomes. HAT positions are classroom-based positions with a reduced teaching allocation to enable them to mentor other teachers, including student teachers, beginning and more experienced teachers, work with university partners and take a role in the school's leadership team. HATs are appointed through a merit selection process which requires, as a prerequisite, application to the NSW Institute of Teachers for consideration of accreditation at Professional Accomplishment or Professional Leadership. These positions are two-year appointments and are limited to 100 positions over the life of the National Partnerships.
Northern Territory	Accomplished Teacher	The NT's Accomplished Teacher status requires applicants to participate in an "inquiry process" over 12 months, based on the NT Teacher Registration Board Accomplished Standards of Professional Practice for Teaching. The assessment of performance is undertaken by assessment panels and moderation committees and includes the appraisal of teaching modelling and role in curriculum and professional learning. This process is currently being reviewed.
Queensland	Experienced Senior Teacher	The Experienced Senior Teacher appointment – subject to an application process – is available to teachers with at least 14 years of experience and 4 or 7 years experience as a Senior Teacher (for 4-year-trained and 3-year-trained teachers, respectively). The classification is offered in recognition of long-service and good practice, and requires an appraisal against a list of criteria. The appointment is not associated with any change in role.
South Australia	Advanced Skills Teacher 2 (AST Level 2)	The position of AST Level 2 is designed to provide recognition for teachers who demonstrate excellent teaching practice. The assessment requires validation of the level of expertise and professionalism and evidence of teacher leadership by a panel consisting of a Site Leader, Equal Opportunity representative and a Peer Evaluator. The assessment involves a portfolio, lesson observation, presentation and discussion. AST Level 2 is currently being reviewed.
Tasmania	Advanced Skills Teacher	The Advanced Skills Teacher position recognises outstanding classroom teachers and leading staff members. It is targeted at teachers recognised as exemplary practitioners, who are accorded additional responsibilities within their school. It is a promotion available to any permanent teacher who satisfies the application process, operating in a similar way to a salary increment. Positions are advertised by individual schools on a need basis.
Victoria	Leading Teacher	The Victorian school system includes one promotional appointment for those teachers who want to remain in the classroom: Leading Teacher. The programme is intended to serve the dual purpose of recognising outstanding classroom teachers; and providing schools with a human resource to lead various in-school programmes and projects. Schools advertise for Leading Teacher positions on a need basis – the position is usually associated with a specific anticipated responsibility. The Victorian Department of Education and Early Childhood Development aims to maintain a Leading Teacher profile of 10 to 15% of full-time teaching staff.
Western Australia	Level 3 Classroom Teacher	Applying for Level 3 Classroom Teacher positions in WA involves a two-stage process. In stage one, applicants submit an application form, introductory statement, written statements and portfolio of evidence, and referees' verification of portfolio statements. The portfolio may include evidence such as students' work, a letter of support from a colleague, extracts from professional learning journal, and items in multimedia format. In stage two, applicants are required to prepare and lead a 45-minute reflective practice session and participate in sessions of other applicants. A reflective practice session includes an oral presentation and facilitated discussion. Applicants are assessed against five teaching competencies that align with Phase 3 of the Western Australian Department of Education's Competency Framework for Teachers. An assessment rubric is used to assess each competency. Each competency is divided into four or five indicators, which must be addressed within each competency. Under the Smarter Schools National Partnership on Improving Teacher Quality, WA has committed to reviewing and expanding the Level 3 Classroom Teacher programme.

Source: Australian Government (2010), Ingvarson *et al.* (2007) and information provided by the Grattan Institute.

Competencies to assess and to use feedback

A range of actors are involved in teacher appraisals. Jurisdictions have varied approaches to the development of competencies to perform an appraisal and to use the feedback from an appraisal. This is summarised below for each of the actors involved in teacher appraisal.

Teachers are the recipients of teacher appraisal but are also actively involved in their appraisal through the self-assessment of their practices for each of the processes of registration, performance management and AST positions. For example, applicants for the AST positions in the Northern Territory reflect on their achievements, strengths and practices against the standards throughout their process. They are supported and challenged by their principal, teaching colleagues and an external panel who consider their propositions and evidence in relation to the Accomplished Teacher Standards. As teacher appraisal processes are generally school-based, the levels of training provided to teachers to support their self-assessment and to benefit from feedback may vary considerably. Information sessions, workshops and training may be available to teachers as well as a range of templates and information sources such as tools for self-assessment.

Principals play the key role in teacher appraisal for each of the processes of registration, performance management and AST positions. In registration processes, authorities commonly provide a range of resources and support measures to ensure that principals can undertake effective appraisals and that staff are supported/guided through the processes. This may include training by teaching statutory authorities. In most schools, the principal is also ultimately responsible for the performance management of the teachers, and may determine what training is offered to line managers/supervisors who are delegated to undertake the appraisals. Formal training in the performance management process is generally available for principals and other school leaders.

Peer evaluators are typical of appraisal for AST positions but may also be used in the cases of performance management and registration processes as part of school-based panels. In the case of AST positions, generally a merit selection process is used to select highly competent teachers for positions as evaluators and an induction course is provided.

External assessors are involved essentially in appraisal associated with AST positions and, in the case of New South Wales, in the monitoring of registration decisions across schools. For example, applications for Level 3 Teacher positions in Western Australia are assessed by two trained assessors. The assessors may be Level 3 Classroom Teachers, heads of learning areas, principals or deputies.

Using evaluation results

The appraisal of teachers against the registration standards forms part of the teacher regulatory system to ensure teachers are qualified, of suitable character and competent to be admitted or to remain in the profession. In a few instances, such as in New South Wales, higher levels of accreditation lead to higher levels of pay. Access to higher levels of registration is also used as a prerequisite to apply to AST positions in some jurisdictions. Finally, even if it is not its primary function, appraisal in the context of registration processes may identify professional development needs to address particular teaching standards.

As part of performance management arrangements, professional development plans for the next cycle are identified to support the teacher's ongoing learning and development. Other objectives include the co-ordination of professional development within the school and informing school policy development, planning and resourcing. In most jurisdictions, the completion of a successful performance management cycle enables teachers to progress to the next increment in their pay structure. In the non-government sector, the results of performance management processes may be used to recognise and reward teachers, including the allocation of performance-based pay.

Applicants who are successful in gaining AST positions receive a pay rise and may assume leadership or mentoring roles in their schools commensurate with their levels of expertise. For instance, in South Australia AST teachers are paid at a higher salary increment and in the Northern Territory Accomplished Teachers receive a 4% pay rise.

Box 4.1 The teaching profession in Australia – Main features

Employment status

Teachers working in the government sector are salaried employees of states and territories. Pay and working conditions are determined by educational authorities of the jurisdictions following negotiations with teacher unions. Teachers working in the non-government sector are salaried employees of schools' organisers which determine pay and working conditions. Most teachers are employed on an ongoing/permanent basis, and this is slightly more common among secondary (81%) than primary teachers (72%) (McKenzie *et al.*, 2008). Teachers access employment on a permanent basis after a probationary period of 12 to 24 months. Teachers appointed to Advanced Teaching Skills positions are typically appointed for a limited tenure of up to 5 years.

Prerequisites to become a teacher and teacher recruitment

To obtain employment as a teacher in Australia, individuals should have a recognised qualification, which is usually a teacher education degree accredited by the relevant teaching authority, or an equivalent foreign qualification. Other requirements include good command of the English language and satisfactory results in a criminal history check.

Teacher recruitment and appointments are typically the responsibility of school leaders and school councils or school boards – both in the government and the non-government sectors – and are undertaken in the context of open competitions. The process may be carried out in consultation with either the educational authority supervising the school or the school's organiser in the case of non-government schools.

Salary and career structure (government sector)

In Australia career progression and salary are almost entirely dependent on length of service and years of initial education. Typically, there is a single salary scale, incremental on the basis of tenure, whose top is reached after 8 to 12 years. Most of the jurisdictions also offer additional classifications which provide access to another salary scale or a salary increment. It is also typical for allowances to be paid to classroom teachers for additional responsibilities such as year level co-ordinator or head of department. Career structures are organised as follows in Australian jurisdictions (Jensen and Reichl, 2011):

Box 4.1 The teaching profession in Australia – Main features *(continued)*

- **Australian Capital Territory**: Single common incremental salary scale with 3 professional phases, each explicitly aligned with years of experience: New Educator; Experienced Teacher 1; and Experienced Teacher 2.
- **New South Wales**: Single common incremental salary scale with 13 steps (a teacher with 4 years of education enters the scale at the 5th step). Additional category of Highly Accomplished Teacher (2-year appointment).
- **Northern Territory**: Single common incremental salary scale whose top is reached after 8 years of service. This is complemented with the category of Accomplished Teacher, a promotion with its own 8-step salary scale.
- **Queensland**: Single 8- or 9-step salary scale for teachers (depending on years of education), complemented with: Senior Teacher Increment (additional salary increment for teachers with a minimum of ten years of service); and Experienced Senior Teacher, a promotion accessible to teachers with several years of experience as Senior Teacher.
- **South Australia**: Single core salary scale for classroom teachers and 2 advanced classroom teacher classifications: Step 9; and Advanced Skills Teacher Level 2.
- **Tasmania**: Single common incremental salary scale with 12 steps complemented with the category of Advanced Skills Teacher, a promotion available to any permanent teacher with its own salary scale.
- **Victoria**: Two classifications: Classroom Teacher with 3 categories, graduate, accomplished, expert, in an 11-step single salary scale; Leading Teacher, with a specific salary scale and additional responsibilities.
- **Western Australia**: Single 9-step salary scale for classroom teachers, complemented with: Senior Teacher Increment (additional salary increment); and Level 3 Classroom Teacher, a promotion with its own salary scale.

Professional development

Professional development for Australian teachers, when regulated, is typically linked to two key professional requirements. The first is the process of performance management which requires the preparation of a professional development plan as part of the annual performance review (this is the case for the Victorian Performance and Development Guide and Queensland's Developing Performance Framework). The second is the participation requirement to obtain the renewal of teacher registration as stipulated by teaching statutory bodies (highly variable across jurisdictions – *e.g.* no mandatory requirement in Victoria; at least 50 hours of professional development activities in previous 5 years, in New South Wales). In some jurisdictions, the educational authority may also define a minimum requirement (*e.g.* 4 to 5 days per year in Tasmania; minimum of 5 days per year in the ACT).

According to a survey conducted among Australian teachers, on average, in the previous 12 months, teachers reported that they spent 9-10 days in professional learning, and leaders spent an average of 12-13 days. Around 60-70% of teachers indicated that the professional learning had increased their skills and capacity to perform their role at the school to a major or moderate extent (McKenzie *et al.*, 2008).

Strengths

Teaching standards have been instituted

Teaching standards, a clear and concise statement or profile of what teachers are expected to know and be able to do, have been established across Australia. They are of two distinct types. First, each jurisdiction's statutory teaching body (except for the Australian Capital Territory) has developed its own set of teaching standards for the registration of teachers and the accreditation of initial teacher education programmes. Second, a number of educational authorities have also developed distinct professional standards for teachers (*e.g.* South Australia, Victoria, Western Australia) – these, where they exist, generally provide the reference for performance management processes and establish the links to the career structure.

Teaching standards are a key element in any teacher appraisal system as they provide the credible reference for making judgments about teacher competence. They strengthen the capacity for educational authorities to effectively assess whether teacher performance meets the needs of school education and whether teachers have attained given levels of competence. A strength in the system has been the extensive involvement of the teaching profession, employers and teacher educators in the development of teaching standards for registration/accreditation. Teaching colleges/institutes as independent statutory bodies provide teachers with professional autonomy and self-regulation and the right to have a say in the further development of their profession. This reinforces the effective use of standards as a lever for the improvement of teaching practices. More challenging aspects to teaching standards in Australia include the multitude of standards (across and within jurisdictions) and their weak linkage to career structures. This will be explored later.

The establishment of national professional standards for teachers is a major development

A particularly significant development has been the creation of AITSL and the ambition to establish a nationally-shared understanding of what counts as accomplished teaching and school leadership. The implementation of National Professional Standards for Teachers by AITSL, planned for early 2011, will provide nationally agreed and consistent requirements and principles to organise the key elements of the teaching profession such as initial teacher education, nationally consistent teacher registration, professional development, teacher appraisal and career advancement. Similarly, the development of the National Professional Standards for Principals by AITSL is intended to support the preparation, development and self-reflections of both aspiring and practising principals to lead 21st century schools across the country. The nationally agreed Teacher Standards will, under Improving Teacher Quality National Partnership arrangements, take priority over the existing standards of jurisdictions.

These national professional standards are also likely to promote the mobility of the teaching workforce and strengthen the alignment of teaching practices to national student learning objectives. They will also serve as a powerful quality assurance mechanism to ensure that Australian teachers and school leaders have the required competencies to be effective educators.

Teacher registration processes are in place

Teacher registration processes are well established in Australian schools (with the exception of the Australian Capital Territory which will implement a permanent arrangement in line with other jurisdictions' processes from 2011). Their main function is that of certifying teachers as fit for the profession mainly through the mandatory process of accessing or maintaining "Full/Competence" status – as such, these processes ensure minimum requirements for teaching are met by practising teachers. Only in the cases of New South Wales, the Northern Territory and Tasmania, are there additional levels of registration accessible to teachers on a voluntary basis. These have the distinct function of guiding teachers' improvement of skills and competencies and steering their aspirations to new roles and responsibilities.

Registration processes constitute a powerful quality assurance mechanism to ensure that every school in Australia is staffed with teachers with suitable qualifications who meet prescribed standards for teaching practice. At their initial level (provisional/graduate registration), they also provide a policy lever for setting entrance criteria for the teaching profession and, through the accreditation of initial teacher education programmes, strengthen the alignment between initial teacher education and the needs of schools. Granting full registration only after a period of employed teaching practice is appropriate as even where there are reasonably high levels of confidence in the quality of initial teacher education, the nature of teaching means that many otherwise well-qualified candidates may struggle to adjust to the demands of the job.

The requirement of registration renewal has clear benefits. It provides incentives for teachers to update their knowledge and skills continuously and it potentially allows the school system to identify core areas in which teachers need to keep improving. Its link to professional development activities also provides the potential to guide the continuing development of practising teachers.

Performance management processes provide a good basis for developmental teacher appraisal

Even if their application varies significantly across schools, teacher appraisal as part of regular employers' performance management processes is expected to take place in Australian schools. In its current form, it has essentially an improvement function with the emphasis on evaluation for teacher development. However, it also performs two additional functions: the identification of underperformance; and the validation of eligibility for salary increment (in most jurisdictions).

The focus on developmental teacher appraisal is suitable. It is intended to identify areas of improvement for individual teachers, and lead to the preparation of individual improvement plans (including professional development) which are supposed to take into account the overall school development plan. Without a link to professional development opportunities, the evaluation process is not sufficient to improve teacher performance, and as a result, often becomes a meaningless exercise that encounters mistrust – or at best apathy – on the part of teachers being evaluated (Danielson, 2001; Milanowski and Kimball, 2003; Margo *et al.*, 2008). Performance management in Australia typically involves helping teachers learn about, reflect on, and improve their practice in the specific school context in which they teach. It generally provides opportunities for teachers to feel engaged which is essential both to gain support from teachers on the appraisal process and to enhance teaching practices.

Advanced Skills Teaching positions grant opportunities for recognition of skills and competencies

Advanced Skills Teaching positions, which exist in almost all educational jurisdictions, for the most part accomplish two important functions: the recognition of advanced teaching skills with a formal position and additional pay; and a better match between teachers' skills and the roles and responsibilities needed in schools through competitions to gain the positions. These have the benefit of rewarding teachers who choose to remain in the classroom rather than move into management positions.

AST positions embody two key concepts in the teaching profession in Australia. First, they recognise the need to introduce career diversification as a result of the greater variety of roles in schools – *e.g.* departmental head, team leader, and manager of curriculum development and/or personnel development. Second, they reflect the need to reward teachers for their developing skills, performance and responsibilities, in what constitutes a competency-based professional career ladder. Teachers, as they access AST positions, are expected to have deeper levels of knowledge, demonstrate more sophisticated and effective teaching, take on responsibility for co-curricular aspects of the school, assist colleagues and so on. Appropriately, access to AST positions involves more formal evaluation processes which are more summative in nature.

Teachers are trusted professionals with a high degree of autonomy and are open to professional feedback

The Review Team formed the view that Australian teachers are generally perceived as trusted professionals among the different stakeholders. This is reflected in the extensive autonomy they benefit in the exercise of their duties. According to a survey of Australian teachers, 74% of primary teachers and 78% of secondary teachers indicated that they were either satisfied or very satisfied with the freedom to decide how to do their job (McKenzie *et al.*, 2008). Teachers are instrumental in contributing to the shaping of their school's strategies to achieve student learning goals. They decide on the teaching content, teaching materials and methods of instruction. Overall, teachers are given considerable scope to exercise their professionalism and benefit from good levels of trust among students, parents, and the communities in general.

There is a good tradition of teamwork in Australian schools. One of the reasons relates to the high degree of teachers' autonomy and the need for teachers to contribute to the school's strategies to achieve student learning goals. Activities such as interpreting and adapting the curriculum to the local context, establishing student assessment methods and ensuring fairness in the grading through extensive moderation processes bring teachers together in activities which stimulate peer learning and increase co-operation within the school.

One of the results of being perceived as trusted professionals is that Australian teachers are generally eager and willing to receive feedback. Teachers generally conveyed to the Review Team that they appreciated the time the school principal took to provide them with feedback and in general found classroom visits, where they occur, useful. Some teachers also revealed being active in seeking feedback from their students. In many cases, the regret was that the extent of professional feedback was limited and they were eager to have more opportunities to discuss their practice.

There is considerable national policy attention to improving teacher quality

Teacher quality is a top policy priority which is reflected, for instance, in the establishment of AITSL and the launch of a major agreement between the Australian government and state and territory governments: the Smarter Schools – Improving Teacher Quality National Partnership. This Partnership entails an investment of AUD 550 million, runs over five years, and involves the financing of initiatives by state and territory governments (often entailing a collaboration among jurisdictions), in the following areas: attracting the best graduates to teaching through additional pathways; improving the quality of initial teacher education in partnership with providers; developing national standards and teacher registration to aid teacher mobility and retention; developing and enhancing the skills and knowledge of teachers and school leaders through their careers; improving retention by rewarding quality teachers and school leaders; and improving the quality and availability of teacher workforce data.

Challenges

Regular teacher appraisal as part of performance management is not systematic across the system and is not perceived as meaningful

There is an expectation in all Australian government schools that teachers go through processes of regular performance appraisal. There is evidence of great variation between schools in the way performance management is carried out, from a very light touch to it through to demanding and elaborate processes in some schools. The Review Team saw examples of principals setting up thorough performance management processes, but also examples of principals who perceived performance management as a simple "signing off" of the teacher's salary increment and the recording of the teacher's needs for professional development. Therefore there are no guarantees in Australian schools that performance management processes are addressing the real issues and complexities of teaching and learning, except in those schools where appraisal is well consolidated.

Appraisal and feedback of teachers seems to take place. According to TALIS,[2] only 14.8% of Australian teachers of lower secondary education reported having received no appraisal and/or feedback from other teachers or members of the school management team about their work (6th lowest figure among 23 countries against a TALIS average of 28.6%). There seems to be, however, less availability on the part of principals to undertake the appraisal. According to TALIS, 30.1% of teachers of lower secondary education reported having received no appraisal and/or feedback from the principal about their work (5th highest figure, against a TALIS average of 22.0%). The Review Team saw examples of principals with little time to perform classroom observation and to engage in a closer analysis of teacher performance.

However, there also seems to be the perception that appraisal and feedback is not meaningful and has little impact. According to TALIS, 63.4% of Australian teachers of lower secondary education agree or strongly agree that in the school the review of teacher's work is largely done to fulfil administrative requirements (the highest figure among the 23 countries surveyed, against a TALIS average of 44.3%). Also, 61.4% of teachers of lower secondary education agree or strongly agree that in the school the review of teacher's work has little impact upon the way teachers teach in the classroom (3rd highest figure, against a TALIS average of 49.8%).

Regular appraisal at the school level is also not perceived as an instrument to reward teachers, which is not surprising as it is not among the main functions of teacher appraisal in the context of performance management processes. For instance, according to TALIS, only 9.2% of teachers of lower secondary education agree or strongly agree that in the school the most effective teachers receive the greatest monetary or non-monetary rewards (4th lowest figure, against a TALIS average of 26.2%). Similarly, only 8.2% of teachers of lower secondary education agree or strongly agree that in the school if they improve the quality of their teaching, they receive increased monetary or non-monetary rewards (4th lowest figure, against a TALIS average of 25.8%).

Teachers have few opportunities for feedback

Australian teachers have relatively few opportunities for professional feedback. The main opportunity to receive feedback on their practices is the annual performance review held with the school principal (or a nominee of the principal). However, school principals are overwhelmed with tasks at the school and, in general they do not seem to have the time to engage properly in the coaching, monitoring, and appraisal of teachers. For example, classroom observations by school principals seem to be relatively occasional. Similarly, the interaction with experts of school review teams is infrequent and does not allow for a comprehensive review of teaching practices for individual teachers.

According to TALIS, the following proportion of Australian teachers of lower secondary education reported that the following were considered with high or moderate importance as a criterion in the appraisal and/or feedback they received: *(i)* Direct appraisal of classroom teaching: 59.9% (5th lowest figure against a TALIS average of 73.5%); *(ii)* Classroom management: 69.8% (4th lowest figure against a TALIS average of 79.7%); *(iii)* Student feedback on the teaching they receive: 58.4% (2nd lowest figure against a TALIS average of 72.8%); and *(iv)* Feedback from parents: 54.7% (2nd lowest figure against a TALIS average of 69.1%). Overall, there is scope for improvement in areas such as classroom observation, peer discussion, coaching, or self-critical analysis.

Teacher appraisal as part of performance management could be more effective in addressing underperformance

There are some indications that teacher appraisal as part of performance management is not effectively fulfilling its function of addressing underperformance. On the one hand, teachers' identified weaknesses seem to be relatively well addressed through support measures provided to teachers. According to TALIS, some support measures to address teachers' weaknesses seem to be more frequent in Australia than in other TALIS countries. The following proportion of Australian lower secondary teachers are in schools where the principal reported that the following measures are always taken to address weaknesses in their teaching as identified by teacher appraisal: *(i)* The principal ensures

that measures to remedy the weakness in their teaching are discussed with the teacher: 65.6% (9th highest figure against a TALIS average of 58.9%); *(ii)* The principal, or others in the school, establishes a development or training plan for the teacher to address the weakness in their teaching: 57.5% (highest figure against a TALIS average of 20.6%); and *(iii)* The principal ensures that the teacher has more frequent appraisals of their work: 20.8% (6th highest figure against a TALIS average of 15.2%).

On the other hand, there seems to be the perception that sustained underperformance is not as well addressed. According to TALIS, only 29.2% of teachers of lower secondary education agree or strongly agree that in the school teachers will be dismissed because of sustained poor performance (11th highest figure, against a TALIS average of 27.9%). Similarly, 42.8% of teachers of lower secondary education agree or strongly agree that in the school the sustained poor performance of a teacher would be tolerated by the rest of the staff (5th highest figure, against a TALIS average of 33.8%). In addition, only 7.1% of teachers of lower secondary education agree or strongly agree that in the school the school principal takes steps to alter the monetary reward of the persistently underperforming teacher (4th lowest figure, against a TALIS average of 23.1%).

The link between the annual performance review and the salary increment does not seem to play any significant function. According to Grattan Institute (2010), previous analysis of teacher evaluation in Australia shows that virtually all teachers receive satisfactory ratings and progress along their career structure so that teacher salaries essentially depend on their tenure (BCG, 2003; Ingvarson *et al.*, 2007). Research conducted by the Boston Consulting Group (BCG, 2003) for the then Victorian Department of Education and Training, referred to in Grattan Institute (2010), estimated that 99.85% of teachers were granted a "satisfactory" outcome on their performance review. In contrast, school principals estimated that up to 30% of teachers were either "below average performers" or "significant under-performers" (BCG, 2003, as cited in Grattan Institute, 2010). It can also be added that the incentive of salary increments linked to performance review does not apply to the majority of teachers which are already at the top of the incremental salary scale (Kleinhenz and Ingvarson, 2004).

Missing links between teacher appraisal, professional development and school development

Even though the necessity of professional development is widely recognised in Australia, the Review Team formed the view that its provision appears not thoroughly planned, fragmented and not systematically linked to teacher appraisal. According to TALIS, only 18.4% of teachers of lower secondary education reported that the appraisal and/or feedback they received directly led to or involved moderate or large changes in a teacher development or training plan to improve their teaching (4th lowest figure, against a TALIS average of 37.4%). Similarly, 16.7% of teachers of lower secondary education reported that the appraisal and/or feedback they received led to a moderate or large change in opportunities for professional development activities (8th lowest figure, against a TALIS average of 23.7%).

There are instances in which the identification of professional development needs is not a requirement of performance management processes. In addition, there is in some cases a lack of clarity about which reference standards are used in annual performance reviews – professional standards developed by educational authorities or standards for registration – to assess teaching performance and professional development needs. This

risks to create inconsistencies between professional development dictated by annual reviews and professional development undertaken to renew teacher's registration.

There is also scope to better link professional development to school development. In our view, school development could better explore its links to the evaluation of teaching practice. This is in part due to the limited time school principals have for pedagogical leadership and the limited extent to which professional development activities are linked to the results of teacher appraisal.

There is little alignment between teaching standards, registration processes and career structures

A problematic aspect of the teaching profession in Australia is that career structures are, in most jurisdictions, dissociated from teaching standards and registration processes. This translates into a detrimental separation between the definition of skills and competencies at different stages of the career (as reflected in teaching standards) and the roles and responsibilities of teachers in schools (as reflected in career structures). For instance, South Australia's Department of Education and Children's Services' set of Professional Standards for teachers provide for four career phases (beginning; established; accomplished; and leaders) which are not linked to the salary structure for South Australian teachers. This is problematic in a range of ways. In particular, it reduces the incentive for teachers to improve their competencies, and weakens the matching between teachers' levels of competence and the tasks which need to be performed in schools to improve student learning.

This challenge is compounded by the fact that, in most jurisdictions, registration standards specify only minimum requirements and do not reflect competencies at different stages of the career (with the exception of three jurisdictions, see below); and career opportunities for effective teachers are limited. There are in general few opportunities for promotion, greater recognition and more responsibility. This is likely to undermine the potentially powerful links between teacher appraisal, professional development and career development.

The extent of externality in teacher appraisal is limited

Teacher appraisal, across its different forms, is mostly school-based and rarely involves agents external to the school. Teacher appraisal as part of performance management processes is organised at the school level and involves essentially its management group; registration processes to access "Full/Competency" status are mostly school-based with little external moderation; and appraisal to gain an AST position is predominantly school-based with some externality in certain cases. According to TALIS, 73.8% of teachers of lower secondary education reported having received no appraisal and/or feedback from an external individual or body (*e.g.* external reviewer) about their work in the school (4th highest figure, against a TALIS average of 50.7%).

The limited extent of externality in teacher appraisal raises a number of challenges. Teachers are appraised according to local interpretations/judgments of common standards with risks of lack of coherence in the application of teaching standards. Teachers are also entirely dependent on local capacity and willingness to benefit from opportunities to improve their practice, see their professional development recognised and gain greater responsibility as they evolve in the profession. The involvement of some externality in

teacher appraisal can provide an element of distance and rigour which can be particularly valuable in validating school-based approaches to teacher appraisal.

The multitude of teaching standards risks sending conflicting messages about teaching

At the end of 2010, a set of national professional standards for teachers was endorsed by all ministers of education. This may address the problematic aspect of the multiple teaching standards in the country and often different sets of standards within an educational jurisdiction. Teaching standards for registration, developed by teaching statutory bodies, and professional standards to guide teachers' careers, developed by educational authorities, co-exist in some jurisdictions. This risks sending conflicting messages about what teachers are expected to know and be able to do at different stages of their careers. Furthermore, it risks weakening the alignment between initial teacher education, teacher registration, teacher appraisal, professional development, and career structure that common reference standards seek to achieve. For instance, there is in some instances a lack of clarity about what standards are used in performance management processes and what standards are guiding the professional development of teachers.

There are some challenges to the implementation of teacher registration processes

There are a number of aspects in the implementation of teacher registration processes which deserve further policy attention. First, the level of externality or external moderation in registration processes might not be adequate – processes are mostly school-based and the interpretation of standards is done at the local level with little moderation across schools. Second, registration standards do not fully reflect the complexity of teaching careers and the different levels of performance achievable with further experience – as only in three jurisdictions are there standards beyond "Full/Competent" status. Third, as the maintenance of registration is essentially based on participation in professional development activities, there seems to be a weak link between registration's renewal and what teachers are actually doing in schools and what their students are learning. Overall, it appears that there is a particular light touch to the renewal of registration. AITSL will be working with all jurisdictions to progress national consistency in teacher registration; this work may address some of the implementation challenges.

A related issue is the consideration by Queensland of the introduction of pre-registration tests in literacy, numeracy and science for all teachers with the objective of "building confidence in teaching standards". This initiative risks sending the message that there is little confidence in initial teacher education providers while also raising questions about the effectiveness of processes to accredit initial teacher education programmes. Also, there is a risk that such initiative will not raise the status and public image of the teaching profession.

AST positions are an incipient approach to career diversification and the reward of teachers

AST positions accomplish two important functions in the teaching profession in Australia: *(i)* Granting teachers opportunities to diversify their careers in response to the roles and tasks performed in schools; and *(ii)* Providing a means to reward teachers for the gained competencies and skills to take on higher responsibilities. While a positive development, AST positions remain, in most jurisdictions, a fairly small initiative to: *(i)* Adequately reach the career diversification which fully reflects today's roles and responsibilities in schools; and *(ii)* Provide proper reward for teachers who excel in their practice.

There are also some challenges to their implementation. In some cases, access to an AST position does not lead to further responsibilities or a change in the teacher's role in the school. This is problematic as the recognition of gained skills and competencies should come along with the ability to take on further responsibilities and the objective of meeting a particular need of the school for those extra competencies. Also, there are instances of programmes where access to AST positions is only allowed to teachers at the top of the basic salary scale. This is also problematic as it does not recognise that teachers can gain skills and competencies at different rates with the risk of demotivating some teachers who might have acquired the skills and competencies to access an AST position before they reach the top of the basic salary scale. Finally, policy needs to address the challenge of incentivising teachers who do not access "advanced skills" status to keep improving.

Policy recommendations

In order to make teacher appraisal more effective in Australia, the Review Team proposes the following approach:

- The alignment of teaching standards with a competency-based career structure;
- Teacher registration conceived as career-progression evaluation;
- Developmental evaluation performed through teacher appraisal as part of performance management, internal to the school, for which the school principal would be held accountable;
- Links between developmental evaluation and career-progression evaluation.

The detailed suggestions are presented below (see Santiago and Benavides, 2009, for a detailed conceptual framework for teacher appraisal). The policy options offer a general set of principles and do not intend to imply that approaches to teacher appraisal across educational jurisdictions and schools should become uniform; on the contrary, the implementation of the more detailed appraisal processes at the state, territory and school level should take into account specific local context and needs.

Align teaching standards with a competency-based career structure for teachers

An important policy objective should be to align the definition of expected skills and competencies at different stages of the career (as reflected in teaching standards) and the tasks and responsibilities of teachers in schools (as reflected in career structures). This would strengthen the incentive for teachers to improve their competencies, and reinforce

the matching between teachers' levels of competence and the tasks which need to be performed in schools to improve student learning. Such alignment can be achieved by developing teaching standards which reflect different levels of the teaching expertise needed in schools; and ensuring levels of teaching expertise match the key stages of the career structure.

A framework of teaching standards is essential as a reference for teacher appraisal. In recognition of the variety of tasks and responsibilities in today's schools and the teaching expertise developed while on the job, teaching standards should express different levels of performance such as graduate teacher, competent teacher, accomplished/established teacher, and leading/expert teacher. These should reflect teachers' tasks in schools and the knowledge and skills that they need to acquire to be effective at the different stages of their careers to achieve student learning objectives. They need to reflect the sophistication and complexity of what effective teachers are expected to know and be able to do; be informed by research; and benefit from the ownership and responsibility of the teaching profession. According to a survey of Australian teachers, 74% of primary teachers and 74% of secondary teachers strongly agreed or agreed that teacher professional standards should be used in any performance appraisal process (McKenzie *et al.*, 2008). The current implementation of National Professional Standards for Teachers by AITSL is a major development in this direction.

The career structure for teachers should then match the different levels of expertise reflected in teaching standards. Such alignment would reflect the principle of rewarding teachers for accomplishing higher levels of expertise through career advancement and would strengthen the linkages between roles and responsibilities in schools (as reflected in career structures) and the levels of expertise needed to perform them (as reflected in teaching standards). A career structure for teachers reflecting different levels of expertise is also likely to enhance the links between teacher appraisal, professional development and career development.

Conceive teacher registration as career-progression evaluation

Given the alignment between teaching standards and the competency-based career structure for teachers, teacher registration can be conceived as career-progression evaluation. Career-progression evaluation would have as its main purposes holding teachers accountable for their practice, determining advancement in the career, and informing the professional development plan of the teacher. This approach would convey the message that reaching high standards of performance is the main road to career advancement in the profession. Teaching registration, based on national-level standards, would be portable across states, territories and school sectors. Access to levels of registration beyond "Competent" level should be through a voluntary application process and teachers should be required to periodically maintain their registration status when not applying to a promotion.

Appraisal for teacher registration, which is more summative in nature, needs to have a stronger component external to the school and more formal processes. It could be a mostly school-based process led by the school principal (or another member of the management group) but it should include an element of externality such as an accredited external evaluator, typically a teacher from another school with expertise in the same area as the teacher being appraised. External evaluators would receive specific training for this function, in particular in standards-based methods for assessing evidence of teacher performance, and would need to be accredited by the proper organisation. It would also

be desirable to establish moderation processes to ensure consistency of school approaches to career-progression evaluation. The reference standards would be the national teaching standards common across all schools but criteria to assess against the standards should account for the school's objectives and context. The main outcome would be the implications for career advancement but it would also inform the teacher's professional development plan.

Appraisal for teacher registration should be firmly rooted in classroom observation as most key aspects of teaching are displayed while teachers interact with their students in the classroom. It should also involve an opportunity for teachers to express their own views about their performance, and reflect on the personal, organisational and institutional factors that had an impact on their teaching. Given the high stakes of career-progression evaluation, decisions must draw on several types of evidence, rely on multiple independent evaluators, and should encompass the full scope of the work of the teacher. Also, student test results as an instrument to assess individuals are challenging. These are not commonly used in countries for the evaluation of individual teachers (OECD, 2005; UNESCO, 2007), in large part because of the wide range of other factors impacting on student results.[3] It remains essential that teachers provide evidence to demonstrate student progress in their classrooms, but it can be provided, for instance, through portfolios or other specific evidence.

Career-progression evaluation is also the basis for recognition and celebration of a teacher's work. It provides opportunities to recognise and reward teaching competence and performance, which is essential to retain effective teachers in schools as well as to make teaching an attractive career choice (OECD, 2005). It does not directly link evaluation results with teacher pay but, instead, to career progression (therefore establishing an indirect link with salaries).[4] This is a desirable option as direct links between teacher performance and pay have produced mixed results, according to the research literature (Harvey-Beavis, 2003; OECD, 2005).

Processes to maintain a given registration status should also be strengthened. In particular, requirements should go beyond professional development activities and include an appreciation of what teachers are actually doing in schools and what their students are learning. This could involve a mostly school-based appraisal of teacher's work based on classroom observation and presentation by the teacher of evidence of good performance. However, there should be an element of externality to registration renewal processes such as the external moderation of school approaches to it.

The Review Team would also not favour pre-registration teacher tests in literacy, numeracy and science as such initiative would not raise the public image of the teaching profession and the status of initial teacher education programmes. A similar objective can be achieved by strengthening selection into initial teacher education, organising diagnostic tests during initial teacher education to identify those in need of support in literacy, numeracy and science, or through more rigorous processes of accreditation of initial teacher education programmes.

Perform developmental evaluation through teacher appraisal as part of performance management processes

The Review Team is of the view that teacher appraisal as part of performance management processes should be conceived as developmental evaluation, *i.e.* the main process through which the improvement function of teacher appraisal is achieved. It

would retain its current character but school-based processes for developmental evaluation would need to be strengthened and validated externally. Given that there are risks of bringing together both the accountability and improvement functions in a single teacher appraisal process, it is recommended that teacher appraisal as part of performance management processes is conceived as predominantly for improvement while teacher appraisal for registration performs a primarily accountability function.[5]

This development evaluation would have as its main purpose the continuous improvement of teaching practices in the school. It would be an internal process carried out by line managers, senior peers, and the school principal (or members of the management group). The reference standards would be the teaching standards but with school-based indicators and criteria. This appraisal should also take account of the school objectives and activity plan. The main outcome would be feedback on teaching performance as well as on the overall contribution to the school which would lead to a plan for professional development. It can be low-key and low-cost, and include self-evaluation, peer evaluation, classroom observation, and structured conversations and regular feedback by the school principal and experienced peers. The key aspect is that it should result on a meaningful report with recommendations for professional development. To be effective, evaluation for improvement requires a culture in which there is developmental classroom observation, professional feedback, peer discussion and coaching opportunities. The willingness to share classroom practice and to receive feedback, which is characteristic of the Australian school system, will surely facilitate this process.

In order to guarantee the systematic and coherent application of developmental evaluation across Australian schools, it would be important to undertake the external validation of the respective school processes. An option is that school review processes, in their monitoring of the quality of teaching and learning, include the audit of the processes in place to organise developmental evaluation, holding the school director accountable as necessary.

Ensure links between developmental evaluation and career-progression evaluation

Developmental evaluation and career-progression evaluation cannot be disconnected from each other. A possible link is that appraisal for teacher registration needs to take into account the qualitative assessments produced through developmental evaluation, including the recommendations made for areas of improvement. Also, in spite of its emphasis on teacher development, teacher appraisal as part of performance management processes should retain its function of identifying sustained underperformance with consequences for both the maintenance of teacher registration and eligibility to salary increment. Similarly, results of teacher registration appraisals should also inform the professional development of individual teachers.

Reinforce the linkages between teacher appraisal, professional development and school development

The linkages between teacher appraisal, professional development and school development need to be reinforced. Teacher appraisal is unlikely to produce effective results if it is not appropriately linked to professional development which, in turn, needs to be associated with school development if the improvement of teaching practices is to

meet schools' needs. The schools that associate the identified individual needs with the school priorities, and that also manage to develop the corresponding professional development activities, are likely to perform well (Ofsted, 2006). Schools can learn from the strengths of effective teachers and implement professional development programmes that respond to their weaknesses.

Effective operation of teacher appraisal and its contribution to school development will depend to a great extent on the pedagogical leadership of school principals. Other education systems have increasingly recognised the importance of school leadership in raising standards, as substantiated in an OECD report (Pont *et al.*, 2008). Teacher appraisal will only succeed in raising educational standards if school principals take direct responsibility for exerting pedagogical leadership and for assuming the quality of education in their schools. Principals are also more likely to provide informal continuing feedback to the teacher throughout the year and not only during the formal appraisal process. More generally, they are essential to make performance improvement a strategic imperative, and help considering teacher appraisal indispensable to teacher and school broader policies (Heneman *et al.*, 2007; Robinson, 2007; Pont *et al.*, 2008). Therefore the recruitment, training, professional development and evaluation of school leaders should be given great importance. In addition, school principals need to spend appropriate time on their pedagogical role. It is our view that the concept of shared leadership needs to be more firmly embedded in schools, to support existing principals and allow them to concentrate on their pedagogical role.

Strengthen competencies for teacher appraisal

An area in which there needs to be particular care is that of the competencies for evaluation. Assessors for teacher registration processes need to be trained to assess teachers according to the limited evidence they gather, the criteria of good teaching and the corresponding levels to attain registration. Assessors should be trained to also provide constructive feedback to the teacher for further practice improvement.[6] Also, substantial activities for professional development on how to best use appraisal processes should be offered to teachers. It is vitally important that teachers are provided with support to understand the evaluation procedures and to benefit from evaluation results. It is also expected that evaluation and feedback become core aspects offered in initial teacher education.

Regarding developmental evaluation, there are advantages to having the principal and/or other teachers as the assessors given their familiarity with the context in which teachers work, their awareness of the school needs and their ability to provide quick and informed feedback to the teacher. However, it might prove difficult for principals to undertake the thorough assessment of each teacher in the school. In addition, most principals have no prior training in evaluation methods and might not have the content expertise relevant to the teaching areas of the teacher being evaluated. Hence, it might prove valuable to build capacity in evaluation methods at the school level by preparing members of the management group or leading/expert teachers to undertake specific evaluation functions within the school.

Notes

1. At the end of 2010, the ACT established a Teacher Quality Institute, which will perform teacher registration functions for all teachers seeking to work in the jurisdiction; the new system is intended to be parallel to other systems. Previously, teachers applying for jobs at ACT schools had to meet requirements set out by the employing body, which were similar to those set in other jurisdictions.

2. OECD's Teaching and Learning International Survey, which was implemented in 2007-08, covering lower secondary education and with the participation of 23 countries (OECD, 2009). The results derived from TALIS are based on self-reports from teachers and principals and therefore represent their opinions, perceptions, beliefs and their accounts of their activities. Further information is available at *www.oecd.org/edu/talis*. TALIS results for Australia are provided in Annex D.

3. In this respect, the development of "value-added" models represents significant progress as they are designed to control for the individual student's previous results, and therefore have the potential to identify the contribution an individual teacher made to a student's achievement. However, in order to be effective, value-added models require vast amounts of data to be collected through large scale national-level student testing across levels of education and subjects, an option with prohibitive costs.

4. According to a survey of Australian teachers, 67% of primary teachers and 70% of secondary teachers strongly agreed or agreed that higher pay for teachers who demonstrate advanced competence would help retain teachers in the profession (McKenzie *et al.*, 2008).

5. Combining both the improvement and accountability functions into a single teacher appraisal process raises difficult challenges. When the evaluation is oriented towards the improvement of practice within schools (developmental evaluation), teachers are typically open to reveal their weaknesses, in the expectation that conveying that information will lead to more effective decisions on developmental needs and training. However, when teachers are confronted with potential consequences of evaluation on their career and salary, the inclination to reveal weak aspects of performance is reduced, *i.e.* the improvement function is jeopardised (see Isoré, 2009).

6. For further details on the range of characteristics and competencies for evaluators see, for example, Santiago *et al.* (2009).

References

Australian Government (2010), *Country Background Report for Australia*, prepared for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, Department of Education, Employment and Workplace Relations, Canberra, available from *www.oecd.org/edu/evaluationpolicy*.

BCG (Boston Consulting Group) (2003), *Schools Strategy Workforce Development*, Melbourne.

Danielson, C. (2001), "New Trends in Teacher Evaluation", *Educational Leadership*, Vol. 58, No. 5, pp. 12-15.

Grattan Institute (2010), *What Teachers Want: Better Teacher Management*, Melbourne, available from *www.grattan.edu.au*.

Harvey-Beavis, O. (2003), "Performance-Based Rewards for Teachers: A Literature Review", paper distributed at the third workshop of participating countries on OECD Activity "Attracting, Developing and Retaining Effective Teachers", 4-5 June, Athens, Greece, available from *www.oecd.org/edu/teacherpolicy*.

Heneman, H., A. Milanowski and S. Kimball (2007), "Teacher Performance Pay: Synthesis of Plans, Research, and Guidelines for Practice", *Consortium for Policy Research in Education (CPRE) Policy Briefs* RB-46.

Ingvarson, L., E. Kleinhenz and J. Wilkinson (2007), *Research on Performance Pay for Teachers*, Australian Council for Education Research.

Isoré, M. (2009), "Teacher Evaluation: Current Practices in OECD Countries and a Literature Review", *OECD Education Working Papers*, No. 23, OECD, Paris, available from *www.oecd.org/edu/workingpapers*.

Jensen, B. and J. Reichl (2011), *Better Teacher Appraisal and Feedback: Improving Performance*, Grattan Institute, Melbourne.

Kleinhenz, E. and L.C. Ingvarson (2004), "Teacher Accountability in Australia: Current Policies and Practices and their Relation to the Improvement of Teaching and Learning", *Research Papers in Education*, 19(1), 31-49.

Margo, J., M. Benton, K. Withers and S. Sodha (2008), *Those Who Can?*, Institute for Public Policy Research (IPPR) Publications.

McKenzie, P., J. Kos, M. Walker, J. Hong and S. Owen (2008), "Staff in Australia's Schools 2007", *Teaching and Learning and Leadership, http://research.acer.edu.au/tll_misc/3*.

Milanowski, A. and S. Kimball (2003), "The Framework-based Teacher Performance Assessment Systems in Cincinnati and Washoe", *CPRE Working Paper Series*, TC-03-07.

OECD (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, OECD, Paris.

OECD (2009), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, OECD, Paris.

Ofsted (Office for Standards in Education) (2006), "The Logical Chain: Continuing Professional Development in Effective Schools", *OFSTED Publications* No. 2639, United Kingdom.

Pont, B., D. Nusche and H. Moorman (2008), *Improving School Leadership, Volume 1: Policy and Practice*, OECD, Paris.

Robinson, V. (2007), "School Leadership and Student Outcomes: Identifying What Works and Why", *ACEL Monograph Series*, No. 41, Australian Council for Educational Leaders.

Santiago, P. and F. Benavides (2009), "Teacher Evaluation: A Conceptual Framework and Examples of Country Practices", paper presented at the OECD-Mexico Workshop "Towards a Teacher Evaluation Framework in Mexico: International Practices, Criteria and Mechanisms", Mexico City, 1-2 December 2009, OECD, Paris, available from *www.oecd.org/edu/evaluationpolicy*.

Santiago, P., D. Roseveare, G. van Amelsvoort, J. Manzi and P. Matthews (2009), *Teacher Evaluation in Portugal: OECD Review*, OECD, Paris, *www.oecd.org/edu/teacherevaluationportugal*.

UNESCO (2007), *Evaluación del Desempeño y Carrera Profesional Docente: Una Panorámica de América y Europa*, Oficina Regional de Educación para América Latina y el Caribe, UNESCO Santiago, 2007.

Chapter 5

School evaluation

Australia has a variety of forms of school evaluation in place, each of which derives from the particular circumstances and traditions of the state, territory and school sector within which it has developed. There are two main forms of evaluation: school self-evaluation and school external performance review. This is represented as a sequence of activities which begins with self-evaluation and proceeds through a planning, reporting and review process which both satisfies external requirements and is an engine of school improvement. External school reviews vary widely across jurisdictions and in government schools work within a clear state or territory policy – typically a School Performance Improvement Framework – and are organised and staffed by relevant state government departments. Particularly positive features of school evaluation include the fact that accountability and transparency are well embedded as national principles guiding school evaluation, the good integration of performance data and survey results into school evaluation processes, the clear rules for school reporting, the recognition of the key role of school self-evaluation, and the existence of well-consolidated external school review processes. Priorities for future policy development include developing a set of national principles and protocols for school evaluation; strengthening the alignment between self-evaluation and external evaluation; defining the nature of externality; ensuring a broad scope for external school evaluation; ensuring a focus on the quality of teaching and learning in both internal and external school evaluation; building expertise among evaluators and improving data handling skills of school agents; and publishing externally validated school evaluation reports to complement the publication of national test data.

This chapter analyses approaches to school evaluation within the overall Australian evaluation and assessment framework. School evaluation refers to the evaluation of individual schools as organisations. This chapter covers both internal school evaluation (*i.e.* school self-evaluation or self-review) and external school evaluation (such as school reviews).

Over at least the last 30 years, school evaluation has become an increasingly significant feature of the educational landscape in countries across the world. Its nature and purpose remains varied, reflecting national traditions, infrastructure and practices, broader educational policy and political agreements. However, there has been a discernable although by no means universal move away from evaluation which emphasises compliance with central policies and procedures towards much greater stress being placed on the need for schools to evaluate themselves as part of wider strategies of school improvement. Partly as a result of this strengthened school autonomy, the role of external agents or agencies has also undergone significant change. Equally, there has been considerable debate about transparency in reporting the results of both external and internal evaluation with concerns about the negative backwash effects of possible "league tables" (based on test results or school review reports and often constructed by the media) being set against the right of stakeholders, particularly parents, to know how well a school is performing, sometimes as part of a wider move towards giving them more choice about which school their child can attend.

School evaluation in an individual country, therefore, must be seen in the context of its particular cultural traditions as well as the wider policy arena if its precise nature and purpose is to be understood. A number of key questions arise when conceiving a school evaluation framework: Where does school evaluation fit within wider system goals and approaches to accountability? Where is school evaluation located on the accountability – improvement spectrum? Is the focus mainly on process or on outcomes? How explicit are evaluation criteria? What constitutes credible evidence – are both qualitative and quantitative measures acceptable? What is the balance between internal and external approaches? What are the expectations about transparency in reporting?

Context and features

Established practices of school evaluation across schools

With its federal constitution, Australia has a variety of forms of school evaluation in place, each of which derives from the particular circumstances and traditions of the state, territory and school sector within which it has developed. However, the principle of school evaluation, whether internal or external, together with expectations about planning and reporting, are established features of the educational landscape across Australia.[1] The recent education reform agenda is giving rise to greater national consistency, particularly in relation to the curriculum, teacher quality and the testing and reporting of school-level data, including performance in literacy and numeracy. However, while approaches to school evaluation have a number of important similarities, its place is less well defined than other elements in the national reform agenda and there remains scope for continued and considerable variation in approach across jurisdictions.

The Australian Country Background Report for this Review (Australian Government, 2010) outlines what it describes as already being typical of the comprehensive nature of school assessment and evaluation throughout the country. This is represented as a

sequence of activities which begins with self-evaluation by the school and proceeds through a planning, reporting and review process which both satisfies external requirements and is an engine of school improvement. The focus is on student outcomes and analysis of strengths and weaknesses leading to the identification of areas for improvement. Evidence is both quantitative and qualitative covering: academic and wider achievement; school culture; student, parent and teacher engagement; teaching quality; and leadership.

There are indications that school evaluations have a significant impact. According to TALIS, the following proportion of Australian teachers of lower secondary education work in schools where the school principal reported that school evaluations (external or self-evaluation) had a high or moderate level of influence on: *(i)* The level of school budget or its distribution within schools: 76.4% (2nd highest figure against a TALIS average of 38.0%); *(ii)* Performance feedback to the school: 96.2% (2nd highest figure against a TALIS average of 81.3%); *(iii)* Performance appraisal of the school management: 88.5% (6th highest figure against a TALIS average of 78.7%); and *(iv)* Assistance provided to teachers to improve their teaching: 86.8% (3rd highest figure against a TALIS average of 70.3%) (see Annex D).

School self-evaluation

School self-evaluation is a feature of all government schools and also features strongly in both the Catholic and Independent sectors. Its precise nature varies across jurisdictions but it is generally seen as contributing directly to developing or monitoring school plans. Planning cycles and formats vary amongst states, territories and school sectors according to local policies and circumstances but there is a general pattern of strategic and operational planning feeding off evidence from school self-reflection. In the most developed approaches, the school self-review gathers and analyses measures of student performance and wider achievements against plans and expectations and has in place a variety of approaches to relate outcomes to inputs, intentions and processes. Helpful frameworks for such reviews are provided in a number of jurisdictions although they vary considerably in the extent to which they make clear the critical points of focus for improvement and the criteria which should be used to judge success. In particular, the need for evaluation to focus on such issues as teacher effectiveness and school leadership, both of which are integral to school success, has not yet been fully recognised.

External school review

Approaches to external school review vary widely across jurisdictions. Generally any external review process in government schools works within a clear state or territory policy and is organised and staffed by relevant state government departments. External reviewers can range from departmental officials to credible individuals with an established track record in running successful schools or with an academic background. Monitoring of the work of schools is typically carried out by local officials who also have some kind of management responsibility for a group of schools (*e.g.* school education directors in NSW, directors of schools in WA). The triggers for more formal or in-depth external reviews can be as a result of specific concerns about performance identified by local officials or perhaps at the request of the school itself.

Use of data

The gathering and analysis of data from assessment and testing together with satisfaction data are established features of most of the various school evaluation systems across jurisdictions. In a number of cases, well-established and sophisticated arrangements are in place which analyse test results in literacy and numeracy across schools in ways which allow comparisons to be made using student-level socio-economic data. Although the approach to test data collection and analysis seemed to be generally seen as helpful, some stakeholders expressed to the Review Team significant reservations about its use for school accountability. Reporting on parent satisfaction is a national requirement of the National Education Agreement and is, for example, integral to the Victorian Accountability and Improvement Framework and schools are required to report on a range of key measures from satisfaction data in their annual reports. Evaluation techniques beyond data collection vary widely with, for example, no clear expectations about direct observation of learning and teaching.

Reporting

Differing approaches to reporting, including accessibility of reports, were formalised in the National Educational Agreement (NEA), together with the Schools Assistance Act 2008, which have created a formal requirement for all schools, government and non-government, to publish an annual report which *inter alia* must include key outcomes and information on satisfaction. Reporting remains a school responsibility with the requirement that reports will be published on school websites and will offer clear and accessible accounts of school performance together with wider reportage about the context and highlights in the school year.

The example of school evaluation in Victoria

Victoria organises its school improvement process in networks of around 20-25 schools, each network being the responsibility of a network leader who reports to the Regional Director. Planning operates on the basis of a four-year strategic plan at both network and school levels allied to one-year implementation plans and annual reporting to the school community. Principals are responsible for undertaking school self-evaluation which drives the planning process. Self-evaluation is expected to focus on the relationship between school practices and student outcomes. Regional network leaders appraise principals against performance targets.

The state Government commissions external organisations to undertake reviews. Reviewers are drawn from former principals, officials or academics and must satisfy criteria covering knowledge of the Victorian education environment, expertise in school improvement and data analysis, interpersonal and communication skills and high ethical standards. They are then subject to an accreditation process and must participate in ongoing professional development. The review process, including the quality of reports, is itself subject to evaluation by stakeholders and officials from the Victorian Department of Education and Early Childhood Development.

External school reviews can be one of four, increasingly intensive, types: negotiated; continuous improvement; diagnostic; and extended diagnostic. The nature of the review in any particular school is dictated by an assessment of risk as indicated by evidence of levels of performance. Reviews are designed to go beyond the conclusions of the self-evaluation process to provide a holistic evaluation of a school's performance and

capacity to improve. They seek to promote internal accountabilities and see the school and the School Council as the main audience.

The example of school evaluation in Western Australia

In Western Australia, self-evaluation is seen as having been a central part of school accountability for at least 20 years. Principals, in collaboration with school staff, make "verifiable" judgements about achievement and about the relationship between school processes and that achievement. Schools publish annual school reports that "describe" the school's performance and compliance with external requirements. The state Government sees the School Report as providing the community with "... a clear sense of how the students are progressing and what is being done to maximise achievement." (Department of Education of Western Australia, 2008).

Directors of schools, in addition to appraising principals and undertaking standards reviews, are expected to maintain regular contact with their schools. Formal school reviews were annual but that was seen as too intense for high-performing schools and a move has been made to vary the length of the cycle based on performance.

Western Australia has an Expert Review Group within the Department of Education and Training which undertakes intensive external reviews of schools based on referrals from Directors. The main focus is on schools giving concern but exemplary practice or specific circumstances can also prompt reviews. A review team undertakes the review and its report includes recommendations for improvement which are followed up six months later. An executive summary is made publicly available.

The national agenda

Against the background of considerable variation in approach across jurisdictions, the developing programme as part of the move to co-operative federalism involving the federal government and the states and territories aims to bring about greater consistency in approach. Agreements reached in COAG and MCEECDYA have provided a clearer framework of national expectations together with new national infrastructure and a firm commitment to improved transparency and accountability (Australian Government, 2008).

The National Education Agreement focuses explicitly and deliberately on outcomes rather than inputs or processes. It makes clear the responsibility of state and territory governments to monitor and review the performance of individual schools in relation to national objectives while recognising the need to take account of local circumstances and priorities. There is an explicit expectation that all schools will meet a common set of high level school performance and reporting requirements. Requirements include participation in annual full cohort national testing of literacy and numeracy and a number of specific requirements relating to forms of national and local reporting. The Schools Assistance Act 2008 applies the same requirements to all non-government schools.

One of the six funded projects within the Smarter Schools National Partnerships, the "School Performance Improvement Frameworks" project, focuses on developing and sharing innovative frameworks for driving improved school performance together with improved understanding of what is needed at system and school level to promote implementation. The approach is being led by Queensland with participation by five other states and territories.

The most visible and controversial element in the reform programme relating to school assessment has been the creation of a website with performance information at the school level. The *My School* website makes results from virtually every school in the country available to parents and the public. Developed by ACARA, the website which went live in January 2010 provides a basic profile of each of nearly 10 000 schools, showing how students at the school performed overall and by band on NAPLAN for each year level tested (Years 3, 5, 7 and 9), compared to the performance of all schools in Australia and compared to schools serving similar students. Based on an index which measures the influence or level of educational advantage that students' family backgrounds have on their educational outcomes at school (the Index of Community Socio-Educational Advantage, ICSEA), comparisons to schools with students from similar backgrounds are intended to make results more meaningful and to enable schools seeking to improve their performance to benefit from the experience of higher performing schools serving statistically similar populations. The similar schools index, however, has been controversial, as the original version was based on community census data rather than the demographics of the actual population of students at a given school. However, the second version of the website uses direct parent data for most schools which provides an increase of 7% to the explanatory power of the index (see below). The website also provides a school statement and basic facts about student demographics, numbers of teachers and school attendance rates. Where relevant to the levels served by the school, data on secondary outcomes is provided about numbers of senior secondary certificates awarded, secondary school completions, Vocational Education and Training certificates, and post school destinations (see Box 5.1).

Through the website, the public and parents can also easily access the results of other local schools as well as those classified as "similar." The ease of information access, intended to promote transparency and accountability, also gave rise to the media creating controversial league tables that rank the performance of all local schools. Although the validity of some of the criticisms is disputed, there were many representations and the threat of industrial action by teachers. In the end, this was avoided with the union calling off the proposed boycott. The then Deputy Prime Minister also asked ACARA to form the *My School* Working Party with representation from the union plus a range of other stakeholders such as principals' organisations and literacy and numeracy specialists. The Working Party was in operation until August 2010 and provided advice on possible enhancements to *My School*.

Box 5.1 Information reported on the *My School* website

School statement:

In this section the school can give an account of the school's mission, values, special programmes, and other information that gives a broader picture of the school.

School facts:

- School sector: government or non-government school
- School type: primary, secondary, combined (primary and secondary) or special purpose (*e.g.* juvenile justice) schools
- Year range offered by the school
- Enrolment: all students (head count) and full-time equivalent enrolments
- Percentage of Indigenous Australian students: Aboriginal and/or Torres Strait Islander descent
- Location: metropolitan, provincial, remote or very remote
- Student attendance rate: aggregated attendance across levels 1-10
- Number of teaching staff: all teachers (head count) and full-time equivalent job load
- Number of non-teaching staff: all non-teaching staff (head count) and full-time equivalent job load

School socio-economic background:

- ICSEA value: The Index of Community Socio-Educational Advantage (ICSEA) is a measure that enables meaningful and fair comparisons to be made across schools. The variables that make up ICSEA include socio-economic characteristics of the area where the students live, the location of the school (regional or remote) and the proportion of Indigenous students enrolled in the school. The average ICSEA value is 1 000 – most schools should have a value between 900-1 100.
- Quarters for each school are displayed in percentages. This gives contextual information about the socio-educational composition of the student population. If students at a school were drawn proportionally from the broad spectrum of the community, then theoretically there would be 25% in each quarter.

NAPLAN results:

- Results are reported as a school average in all tested subjects
- Results are compared to schools with students from similar backgrounds and all Australian schools
- Participation, absentee, exemption and withdrawn rates are reported: school and national average
- Indicative confidence intervals for the results

Senior secondary outcomes (data are not comparable between jurisdictions):

- Number of seniors who have completed secondary school
- Number of seniors who have completed a specific training programme (*e.g.* VET)
- The post-school destination of former seniors (vocational study, university study or in employment)

Source: Reproduced from Rosenkvist (2010).

Subsequently to the visit by the Review Team, version 2.0 of the *My School* website was launched in March 2011. Among the new features, the site provides access to school financial information, with directly comparable details of recurrent income and capital expenditure for all government and non-government schools. Another important feature is the addition of a third year of results for NAPLAN allowing users to now follow trends in school performance over time. Other improvements include a new method of calculating each school's value on the Index of Community Socio-Educational Advantage (ICSEA) by using data supplied by parents and including factors such as the proportion of students at a school from a language background other than English. Improvements to data presentation that make it simpler to compare statistically similar schools have also been made.

The non-government sector

While the requirements relating to the *My School* initiative apply across government and non-government sectors, school evaluation practices in the Catholic and Independent systems are not mandatory but must reflect the national framework and goals for education. Catholic schools operate within system requirements which emanate from church authorities and, while arrangements differ across the country, clear and consistent common features are evident. The Sydney Diocese in New South Wales, for example, has a set of indicators "How Effective are our Catholic Schools", adapted from the framework for inspection and self-evaluation used in Scotland (HMIE, 2010). The Framework focuses on teaching and student outcomes, Catholic identity and stewardship. Schools are expected to self-evaluate against the framework and are subject to external challenge from officers from the Diocese. Schools are required to make reports publicly available annually covering student examination and testing data.

Independent schools are more self-contained but must satisfy a range of stakeholders, including government, and are reported to have generally in place effective mechanisms to report on progress towards stated goals and clear targets for improvement. Registration requirements vary across the country and, as in Western Australia, can include formal school reviews by the registration authority. In Victoria, the Victorian Registration and Qualifications Authority (VRQA) monitors the performance of Independent schools against minimum standards which are prescribed for schools in all systems.

Move towards greater rigour and transparency

It is not possible in this brief overview to do justice to the diversity and complexity of arrangements which are made for school evaluation and reporting across jurisdictions and systems. What is clear is that there is a general move towards greater rigour and transparency with a strong focus on student performance in literacy and numeracy. In particular, developments in national testing and reporting are having a very powerful effect on thinking about school evaluation across all stakeholders.

Strengths

Accountability and transparency are well embedded as national principles guiding school evaluation

The developing culture of school evaluation and improvement across Australia has already become particularly well established in a number of jurisdictions. The national

policy environment has transparency and accountability as key planks in its improvement agenda. The language of accountability and transparency at the national, system and school levels is well aligned. School self-evaluation is an expectation and some form of external review mechanism is increasingly common. Test results, focusing on literacy and numeracy, are widely used to inform evaluation.[2] Annual reporting at school level is a requirement and the recent creation of the *My School* website has reinforced the policy commitment to create open and transparent benchmark data. In addition, results of school evaluation are widely publicised: according to TALIS, 75.7% of Australian teachers of lower secondary education are in schools where school evaluation results (external or self-evaluation) were published (4th highest figure among the 23 TALIS countries, against an average of 55.3%, see Annex D).

The nature of the federal system has historically meant that there are wide variations in approach but the emerging consistency of policy statements at all levels is impressive and the need for evidence-based evaluation to drive school improvement does not seem to be in question. The national policy statement on the educational revolution, for example, makes it very clear that accountability and transparency are integral to the overall strategy (Australian Government, 2008). On page 31, it states:

> *Clear accountability helps create a learning environment that encourages innovation and excellence from school leaders, teachers and students. It also means that students, parents and teachers have the evidence they need to make informed choices.*

That commitment is reflected in the policy agendas of states, territories and in the Catholic and Independent systems and was echoed in discussions undertaken with officials, principals and teachers at both system and school levels during the Review. It is also worth noting that the articulation of policy is not characterised by assertion but is often supported by clear references to sources of evidence which have influenced the direction of travel.

The use of performance data and survey results is well integrated into school evaluation processes

A striking feature of our discussions about school evaluation was the extent to which the need to have valid and reliable data was rarely questioned. Sophisticated forms of testing and data analysis have been in place for some time, notably the SMART approach in New South Wales (see Box 2.2 in Chapter 2). Such data not only provide teachers with valuable diagnostic evidence about young people's performance but also help to identify issues in relation to learning and teaching and the performance of the school more generally. The use in a number of jurisdictions of relatively fine grained socio-economic data at the level of the individual student was helping to build confidence in the robustness of comparative data. The emphasis on student outcome data also helped to guard against subjective judgements dominating decision making.

Quantification is not confined to test results. Stakeholder surveys are already an established feature of school evaluation in a number of jurisdictions and are now a requirement of school reporting. While particular instruments are not always mandatory, the principle of gathering evidence about perceptions and levels of satisfaction is now an expectation and examples of effective practice are increasingly evident.

The principle of publishing performance data at the school level is established

The move to publish results of NAPLAN testing on a school-by-school basis remains controversial. Educational change can often proceed at a very measured pace, seeking to build wide consensus before action. However, the NAPLAN and associated *My School* website represent a powerful example of how a clear and well-articulated policy allied to determined and consistent leadership can bring about quick change. The creation of the *My School* website has challenged sections of the educational community and, by a combination of clarity and flexibility, the nature of the debate seems to have moved from questions about the principle of publication to more specific issues to do with the content and form of presentation. Clarity of purpose has sent a powerful message that the principle of publishing test results with comparative school performance was not in question. Flexibility in agreeing to modify the index of socio-economic advantage, improve the sophistication of the analysis including a form of value added (originally planned but which requires data for one cohort to be available for two years, *e.g.* 2008 and 2010), and to extend the data on schools (such as financial information, included in version 2.0 of *My School* in March 2011) have contributed to bring consensus to the principle of publishing performance data at the school level among stakeholders.

The key role of school self-evaluation is recognised

The strong emphasis on self-evaluation is a clear strength of the approach. Principals and school leadership teams have the responsibility for gathering and analysing evidence about current performance against expectations. In this way, planning cycles are typically built around self-evaluation/reflection, using quantitative and qualitative data as evidence for decisions about priorities in improvement planning and as part of school accountability.

There are clear rules for school reporting

The recent agreement that all schools must publish annual reports on their websites is an important development in transparency and accountability. Following the Melbourne Declaration on Educational Goals for Young Australians in December 2008 ministers agreed that public reporting on Australian schools would: support improving performance and school outcomes; be both locally and nationally relevant; and be timely, consistent and comparable. In June 2009, they agreed a set of eight principles and related protocols for reporting on schooling in Australia, the Principles and Protocols for Reporting on Schooling in Australia (MCEECDYA, 2009). This is a powerful document which makes clear their commitment to transparent accountability. The principles relate directly to data on student outcomes and information about the school context and resourcing. The protocols are designed to promote the integrity of the process and to provide safeguards against simplistic comparisons being made amongst schools.

External school reviews are well established

Some form of external review is widely in place across jurisdictions. The nature of externality is very much a matter for the jurisdiction concerned but the need for a view from outside the school itself seems to be common practice. A number of jurisdictions have recognised the need to engage reviewers who do not have any direct responsibilities associated with the school. The use of successful principals, serving or retired, is one such approach. In Victoria, for example, organisations are contracted to provide school reviews and reviewers are required to satisfy criteria relating to their skills and experience

before being formally accredited. In Western Australia, regular reviews are conducted by departmental officials who have management responsibilities in relation to the school but the more in-depth reviews are conducted by an external review team working directly to the Department's Expert Review Group. In Queensland, the new Learning and Teaching Audits are independent of the school by a team drawn from 16 highly successful principals along with executive directors from outwith the line management of the school being audited. In New South Wales, review teams always include an external member and for those schools deemed to be at greatest risk of failure, the extent of externality is increased.

The frequency of external reviews again varies significantly across jurisdictions. There does appear to be a move towards a risk-based determination, using available data to allocate schools to categories of risk which in turn determine the frequency, depth and degree of externality of reviews. Variants of this approach are evident in Western Australia, New South Wales and Victoria. By contrast Queensland sees its new audit system as a tool to assist all state schools as part of continuous improvement. More generally there is a relationship between regular (at least annual) external monitoring normally undertaken by a departmental official and periodic external reviews on cycles which depend on a determination of risk.

The framework for reviews and the criteria used to inform judgements again vary across jurisdictions. The detailed set of guidelines and indicators, "How Effective is our Catholic School", provides a developed example of a comprehensive approach to such a framework. Judgements are made on a seven-point scale against performance indicators and data is gathered and analysed using the SMART package developed for government schools in New South Wales. Learning and Teaching Audits in Queensland cover defined aspects of a school's work: improvement; data on performance; culture of learning; use of resources; teacher expertise; teaching and learning. In Victoria, the Association of Independent Schools of Victoria has developed a benchmarking tool which brings survey and performance results together and which it recommends for use by governing bodies.

National initiatives promote innovation in approaches to school evaluation

The national approach to promoting innovation through centrally funded projects within National Partnerships is an important element in the developing national agenda. The Smarter Schools National Partnership "School Performance Improvement Frameworks" project described earlier in this chapter is at a very early stage but it has the potential to deepen understanding about school improvement and the place of evaluation in building capacity. In particular, the developments in Queensland have clear strengths which are likely to have a strong impact on developing thinking about school improvement across Australia.

Registration processes are in place and well integrated in school evaluation frameworks

The registration process for non-government schools also employs forms of external review. Registration is undertaken at the state or territory level but, for the Catholic system, is devolved to the church authorities. In the Independent system, the schools themselves have a direct interest in being able to demonstrate their effectiveness and often employ external reviewers as part of their own development processes. Registration varies considerably in its rigour but, in Western Australia for example, the registration

cycle can be up to seven years with a detailed "inspection" at the end of that period. The Catholic sector has its own registration process involving monitoring visits. The Board of Studies in NSW employs 12 inspectors who must have a teaching qualification and a successful track record. They visit schools accompanied by a Board official for up to two days at least once every five years. More frequent visits are made to new schools or those giving concern. Aspects covered include learning and teaching, the quality of lesson preparation, record keeping, policy framework, and child protection. The report is not public but the school will usually provide a summary in its annual report.

Challenges

Developing national consistency while allowing legitimate diversity

Given existing wide variations in policy and practice in relation to school review across Australia, a major challenge lies in determining what constitutes a desirable measure of consistency as against legitimate diversity. The nature of the emerging national agenda is likely to be strengthened by more effective school evaluation which has sufficient consistency across jurisdictions to allow it to have national credibility and to extend information about school performance in ways which complement existing test results. Transparency at the national level through the presentation of test results on the *My School* website would be enhanced by the kind of evidence and evaluation which credible school evaluation can provide.

It is clear that much of what is required is in place in aspects of current practice across jurisdictions and school sectors. The challenge is to articulate a national strategy for school evaluation which builds on the best of current practice and continues to allow flexibility of approach within agreed parameters. School review should not be seen as a threat or as something which only applies to situations giving rise to concern. Just as the approach to testing and the presentation of data had been developed in ways which stress its utility and which seek to minimise perverse effects, so there is a need to use school review as a key driver for improvement in all schools. It can improve accountability by ensuring that evaluation relates to the wider agenda of educational quality outlined in the Melbourne Declaration.

An important ingredient of the debate is an accepted model of school effectiveness. Such broad model would provide clear criteria for effective schools and provide a robust, research-based foundation for school evaluation. Hence educational jurisdictions would benefit from a coherent overall framework for school evaluation drawing on a rationale for school effectiveness.

There is little national direction on the role and nature of school reviews

National and local policy statements stress the importance of accountability and transparency but the outworking of these principles tends to focus almost exclusively on data and information. School review and reporting are accepted features of the overall strategy but there remains a need to clarify a number of vital issues relating to the relationship between the role of reviews in both accountability and improvement; the scope of reviews in relation to the emerging national agenda; the critical areas on which reviews should focus; the role and nature of externality; and the extent of transparency.

Different jurisdictions have addressed mixtures of these issues in their own context but no clear national direction of travel has as yet emerged.

Too great reliance on measuring and publicising student outcomes can have undesired effects

A key plank in the national policy agenda is the belief that measuring and publicising student outcomes on a comparative basis will lead schools to focus on taking the action necessary to improve their relative performance. Thus increased accountability and transparency will help drive improvement. There are, however, a number of possible perverse effects in placing too great reliance on this approach, not least the risk of a possible narrowing effect on the curriculum and wider achievement with an overemphasis on that which is assessed through the tests. There is also a danger that schools which perform satisfactorily may become complacent as the spotlight falls on those schools which perform least well comparatively. During the Review visit, teacher, principal and parent representatives raised instances and anecdotes of the perception that the high stakes of NAPLAN results could lead to:[3]

- Curriculum narrowing, when teachers and schools focused on what was tested;
- Time diverted from regular curriculum for special test preparation for NAPLAN;
- Concentration of resources on students just below minimum proficiency standards and inattention to the lowest performing students in order to maximise the number of students scoring proficient;
- Asking Indigenous students and low performing students to stay at home on test days so as to increase school test performance;
- Negative effects on teacher-based assessments and student engagement in rich curriculum tasks through which teachers can genuinely understand student learning.

Stakeholders also were consistent in their concerns about the limitations and potential adverse consequences of reporting on the *My School* website. Stakeholder groups, including representatives of the Australian Government and of states and territories, were uniformly concerned about league tables that had been constructed by the media from information easily available on the website: they agreed that league tables are misleading, inappropriate and should be discouraged for any number of reasons, *e.g.* school outcomes are affected by a number of factors outside the control of schools, raw results tend to reflect socio-economic status, different measures of school results lead to different rankings of schools, effects of mobility, etc. They also raised concerns about the public debasement of schools with low results, the potential of labelling of schools and students as failures and of reinforcing stereotypes about certain subpopulations, and the perverse incentives provided for schools to rig their results. Examples of "rigging" include the possibility of schools selecting students with an eye toward their test scores rather than equity, in addition to the potential negative effects noted just above. In addition, some observed that the implicit competition encouraged by the *My School* website (*i.e.* to be labelled as higher performing than similar schools) may discourage collaboration between schools although the same information could also promote learning from high-performing schools which have similar characteristics.

The challenge, therefore, is to harness the power of publication of the quantitative data by ensuring that it is "felt fair" by the school concerned and that it is set in a wider set of evidence about performance which reflects the wider agenda set for Australian education in the Melbourne Declaration. That implies the development of a wider strategy which uses school evaluation evidence in ways which encourage schools to remain aspirational in relation to the wider educational agenda whatever their test results. As the new Australian Curriculum becomes embedded in schools, a major challenge will lie in ensuring that the full scope of its expectations is realised and that sufficient attention is given to raising performance across the areas it covers.

There is a need to improve the scope of the information provided by the My School website

There is a concern that by giving primacy to NAPLAN, the *My School* website provides and encourages a very limited view of the skills and knowledge students need to lead productive and rewarding lives and, as noted above, a narrow view of the goals of schooling. Similarly, there is concern that, in aiming at strict national comparability across schools, the limited *My School* data was isolated from other available data in some states and territories about local schools that could provide a richer picture of school quality and student learning. For example, in Victoria, school reports routinely include A-E scores, NAPLAN performance, Senior Certificate results, and climate results from parent and teacher surveys.

Given the importance of the *My School* website as a central part of the accountability and transparency agenda, it will be important to ensure that the content of the site continues to develop in ways which improve its utility and acceptability. The improvement in the socio-economic weighting by moving to individual student data together with the capacity to provide value added measures are both important steps forward. However, many of the perceived difficulties associated with the existing approach could be mitigated by also providing better access to wider evaluation evidence of the kind contained in credible reports of school reviews. The requirement for all schools to report annually is an important part in that process. The challenge remains to ensure that annual reports convey straightforward messages about school performance to complement the statistical data on *My School* in ways which command the confidence of stakeholders in relation to its objectivity and openness. If there were clear expectations about the criteria used for evaluation and the role of external confirmation, then parents, politicians, officials and the wider community would all have access to a more holistic view of the school's performance in forms which allowed comparison and benchmarking. In essence they would have an authenticated narrative not just numbers to help form a view about the quality of a school. However, to be fully meaningful to all stakeholders, that narrative must be expressed in ways which convey clear and simple messages and do not require highly sophisticated understanding of either statistics or education.

It should be noted that new indicators are now available on version 2.0 of the *My School* website (released in March 2011): growth in student performance (for those students who sat the 2008 and 2010 NAPLAN tests at the same school and have results at two year levels); school recurrent income and capital expenditure information; and proportion of students with language background other than English. There are also plans for future versions of the *My School* website to include information on student, parent and teacher satisfaction; proportion of students with a disability and numbers/proportions of teachers by level of expertise under new national scheme.

Clarity is needed about the nature of externality

The relative contributions of self-evaluation and external evaluation also need to be set against the overall purposes of school review. Self-evaluation has the merit of being immediate, responsive to the school's particular needs and circumstances and its results are "owned" by the school. However, self-evaluation which serves the needs of accountability is subject to inevitable tensions between rigour and depth on the one hand and a natural desire not to undermine the confidence of parents and superiors on the other. Even if the results of self-evaluation are seen a wholly a matter for the school, internal politics and power relationships will still influence the rigour of the process. There is also a limiting effect arising from understandable reluctance on the part of those who are strongly committed to a particular course of action to recognise or accept negative evidence. Such limitations suggest that self-evaluation is more a tool for managing development than for challenging assumptions or for arriving at conclusions which threaten key actors in the school's hierarchy. There will always remain issues of credibility amongst stakeholders in accepting that the story being told by those who are accountable for success is dispassionate and accurate.

The involvement of externality in school review, therefore, both provides that element of distance from the internal dynamics of the school and gives the kind of perspective and challenge to assumptions and to the interpretation of evidence which can lead to greater rigour in the process. Credible externality lends authenticity to the outcomes of evaluation. However, externality can be achieved in a variety of ways; who evaluates, what is evaluated and how, and the ways in which the results are agreed and communicated must be explicit from the outset. Common practice across Australia is for officials from within a particular jurisdiction to be the external element in some reviews. Of course, those officials are themselves part of the managed structure within which the school operates and are therefore subject to the same constraints about relationships, authority and consequences which apply within the school itself. In a number of other cases, external reviews are triggered by assessments of risk or by referrals by officials. While these approaches do provide greater distance, the approach reinforces the view that externality is somehow associated with failure rather than being a necessary element in evaluation irrespective of prior assumptions about a school's performance. Clarity is needed about the nature of externality and about the contexts within which it is important. When confined to the most negative cases, the danger is that school review becomes something which is done to a school and is a "badge of failure" rather than an important element in the improvement and reporting process for all schools. An interesting development is the Learning and Teaching audit process in Queensland which will apply to all schools, has a team of specially recruited and highly credible evaluators and has a clear focus on learning and teaching.

The focus of school review needs to be better defined

Clarity about the focus of school review is also important. Reviews need both to evaluate the outcomes being achieved and to identify the key factors which have influenced those outcomes. Reviews need to take direct account of those factors which are central to school improvement. Those factors include the quality of teachers and the teaching process; the ethos of the school; leadership; and the capacity of the school to evaluate itself. It is important, therefore, to have a framework of criteria for evaluation which requires evidence about each of these factors and their relationship to the school's performance. There are strong examples of aspects of this approach in different states and

territories and in both the Catholic and Independent systems. However, there remains considerable variation in approach. Leadership in particular, a key factor in school effectiveness, does not seem to figure strongly in evaluation frameworks. The Smarter Schools National Partnership "School Leadership Development Strategies" project being undertaken by AITSL aims to develop a national approach to enhance school leadership capacity.

The degree of follow-up to school reviews is variable

The feedback schools receive from an external review is a major input into school improvement processes. Typically a school review is followed by the formulation of recommendations for improvement which the school is supposed to implement following the preparation of an improvement plan. However, the Review Team formed the view that the degree of follow-up by school review authorities was variable, including within an educational jurisdiction, depending on the capacity of regional networks. Without evidence-based feedback and mechanisms for monitoring and following up subsequent action, school reviews may have more limited impact on school improvement.

There is a need to build capacity for undertaking evaluations and using their results

Of course, the quality of any evaluation process is highly dependent on the capabilities of those undertaking the evaluations and on the ability of users to interpret results. The *My School* approach with its colour coding system has the clear benefit of simplicity. The quality of the analysis is promoted by the ability to confine test construction and analysis to experts in that field. If that analysis is taken on trust then users can move quickly to considering implications and actions. Any move to a more inclusive and holistic approach may jeopardise the quality of the evidence due to lack of expertise on a variety of fronts. In particular, a stronger focus on the quality of learning and teaching in classrooms requires an evaluator to have more than personal competence as a teacher or school manager. Credible evidence from classroom observation requires particular skills relating to observing and recording the essentials of teacher-pupil interaction. Similarly, approaches to testing the reliability of evidence through triangulation and other forms of cross-referencing have not been developed systematically across the country. However, there are a number of examples from across Australia of ways in which capacity building is being addressed. While there are examples of evaluators lacking credibility, attention is being paid to the need to select the right people and to give them additional training. Good principals and good teachers do not automatically make good evaluators.

Policy recommendations

Develop a set of national principles and protocols for school evaluation

The challenges identified in the previous section give rise to a number of important policy considerations. The ultimate test of the strong themes in the educational reform agenda in Australia relating to curriculum, teacher quality, transparency and accountability will be their positive impact on schools and classrooms, ultimately leading to improved student learning. School evaluation has the potential to help bring coherence

to the wider agenda, provide the kind of formative evidence which can inform both policy and practice and improve traction at the point where formal learning is taking place. Realising that potential requires agreement about the nature and purpose of school evaluation within the overall reform strategy together with the clear communication of expectations and the development of approaches which command the confidence of all stakeholders.

The federal constitution of Australia requires that any programme of development must be established through the kind of consensus building which has characterised other aspects of the reform agenda. The current strengths which are evident across jurisdictions must be respected and built upon. Along the lines of what was proposed in Chapter 2, an important first step might be to agree across jurisdictions a set of principles and protocols for school evaluation along the lines of that produced for reporting in June 2009. Such a policy statement should address the issues identified in this report in ways which build on current best practice, align with the policy agenda and respect traditions of Australian schooling.

The proposed set of principles and protocols would need to address a number of important issues identified elsewhere in this chapter. The first is to be clear about the degree of national consistency which is desirable. School evaluation in Australia takes different forms and serves different purposes across jurisdictions and school sectors. Although pilot studies are being taken forward as part of the National Partnership programme, there remains a key question about how far school review and evaluation should become a more central plank of the educational reform agenda. In particular, there is scope to use a more consistent and robust approach to school evaluation as a means of relating accountability and improvement more directly to the broad goals for Australian education set in the Melbourne Declaration and to improve the impact of the suite of reforms at the school and classroom levels. More consistency in the nature and form of school evaluation would also make a significant contribution to policy formulation at the state and national levels. The insights into the reasons for patterns of performance which could be distilled from aggregations of school reports could provide guidance about the extent to which improvement was needed in teaching, resources, leadership, etc. Overview reports of that kind would allow a more informed alignment to be achieved between the national policy agenda and the reality of school performance in the round.

Clearly establish the fundamental purpose of school evaluation

As part of a general agenda, the fundamental purpose of external evaluation needs to be more clearly and consistently understood. School evaluation can be part of the strategy to bring about general improvement across all schools or, more narrowly, it can focus on “failing schools”. The approach adopted depends on the underlying policy agenda and the evidence about the performance of the school system as a whole. However, a rigorous but constructive approach to evaluation is seen by many countries as a means of driving improvement while also satisfying the needs of accountability.

Strengthen the alignment between self-evaluation and external evaluation

Moves towards achieving a much closer alignment between self-evaluation and external evaluation are evident across a number of European countries. All four countries in the United Kingdom, for example, have variants of such an approach with clearly established frameworks which encompass both internal and external evaluation. In

England, the approach combines school evaluation with an extensive programme of national testing which is similar to that which has been developed in Australia. In Scotland, there is a framework of quality indicators (How Good is Our School) which is designed to cover good practice across the key influences on school quality and to relate these not only to student attainment in formal tests and exams but also to wider student achievement in line with the Scottish programme of curriculum reform, Curriculum for Excellence (CfE). CfE has many similarities to the new Australian Curriculum. The central requirement is that internal evaluation and external evaluation use common criteria and share a common language of quality. Where this is not the case, the school can be pulled in a variety of different directions with no strong evidence base to determine priorities. The criteria can be expressed in different ways but they should focus on those areas which are known to be critical factors in school quality. An example of such an approach can be seen in "How Good are our Catholic Schools" which is already in operation in Australia.

Define the nature of externality

Another important policy question relates to the nature of external evaluation itself. Who are suitable external evaluators and what should be their relationship to the school and to those who manage the school at the national, state, territory or school sector levels? Externality implies sufficient distance from responsibility for the school's performance to avoid conflicts of interest and perceived bias. Where officials of the authority are used then safeguards must be built in to address these independence issues. The use of individuals who do not owe their allegiance to the jurisdiction concerned is clearly desirable as seen in Queensland, Victoria and other jurisdictions. There is no single, prevailing approach to who should be engaged in external evaluation but there is a need to establish clear expectations about externality which will apply across jurisdictions.

Experience internationally also provides a range of models of external evaluation bodies or mechanisms including well-established inspectorates and review bodies as in the UK countries, Ireland, the Netherlands, Belgium, France, Portugal, the Czech Republic, New Zealand etc. Similarly, in countries like Sweden there has been a recent move to establish a strong inspection system which relates to all schools and which also builds on self-evaluation. Elsewhere in Europe, while there are no central inspectorates, more specialised evaluation teams have been established as seen in some of the German *Länder* and Denmark. In most cases, there are clear safeguards to ensure independence from the school or local authority being inspected, usually by direct reporting at the national or state level. However, inspectorates are only one mechanism for creating such independence and arrangements such as those in Victoria or New South Wales where local managers are on teams but are complemented by external team members can also be made credible with appropriate safeguards.

Ensure a broad scope for external school evaluation and place greater emphasis on follow-up

The scope and frequency of external review are also important issues. Moves in the latter part of the last century to extend the scope of inspection in many countries, but perhaps most notably in the UK, have now given rise to serious concerns about over inspection and have led to moves towards more risk-based and proportionate approaches. In the Netherlands, for example, inspection has moved significantly in recent years

towards a risk-based approach which concentrates on those schools which are identified as significantly underperforming. In Scotland, inspection has become more directly focused on improvement by emphasising the importance of self-evaluation and using inspection to validate a school's approach to self-evaluation. In Australia, while national policy has a particular focus on literacy and numeracy, the Melbourne Declaration and the creation of a broad Australian Curriculum points towards a broadly-based improvement strategy. Indeed, literacy and numeracy are significant because they are major contributory factors in the relatively high proportion of young people who are not succeeding at school. That issue is not necessarily concentrated in particular schools but requires all schools to be addressing underachievement. At the same time implementation of the broadening Australian Curriculum, which again applies to all schools, also suggests a more general focus than that which a "failing schools" agenda might imply. For these reasons, developing policy on school evaluation in Australia should seek to use its potential to challenge complacency and provide evidence about progress on a broad front.

The external evaluation of schools has typically the advantage of granting the possibility to compare performance across schools and to assess performance against reference standards. However, external evaluation also runs the risk of focussing on commonalities rather than uniqueness in its attempt to seek comparability and generalisation (Nevo, 2002). Such approach might overlook the local perspective and special needs of the school. Hence, arrangements to external school evaluation need to include strategies to account for local perspectives, context, needs and constraints. This reinforces the need for a close articulation with school self-evaluation.

If school reviews are to have an impact on school improvement, follow-up by school review authorities need to become more systematic and resourced with the objective of supporting schools in the implementation of their improvement plans. It seems as if school review processes are already producing a relatively high amount of feedback while further investment needs to be directed at strategies to ensure that schools effectively use the feedback they receive. The extent of follow-through activities by school review authorities could be made dependent on the extent of improvement needed by a school and its capacity to improve.

Ensure a focus on the quality of teaching and learning in both internal and external school evaluation

It is important that school evaluations do not focus simply on the relationship between policy, planning and outcomes. The most important contribution which school evaluation can and should make to understanding the performance of a school is its focus on learning and teaching. The quality of teaching is central to the quality of young people's learning and the key variable which a school can influence. The central task of school evaluation, therefore, is to determine the quality of teaching across the staff as a whole. This can be a sensitive issue but sends the signal to pupils, teachers and parents that school evaluation is not a bureaucratic exercise which is largely the concern of school managers but relates to the work of each and every member of staff. Wider evidence about compliance with expected procedures and student outcomes can then be interpreted in ways which promote action at the classroom level. School evaluation looks "inside the black box".

Build expertise among evaluators and improve data handling skills of school agents

Similarly, the skills and expertise of evaluators are important. Knowledge of education and a strong track record of success in the field are probably necessary but not sufficient conditions for such evaluators. Interpersonal skills and the ability to communicate clearly and succinctly are both vital as is the ability to absorb, analyse and synthesise complex evidence including statistical data. Again a number of the approaches used in Australia are very sensitive to the needs of evaluators and there are good examples on which an elaboration of the approach can build.

Of course, given the emphasis on self-evaluation, the evaluators are not confined to specialists in this field. There is a need to ensure that all of those who must gather evidence and analyse results have the necessary skills in class observation, interviewing, data gathering, analysis and interpretation which both ensure validity and reliability in the evaluation process and which allow the results of evaluation to be understood and translated into action. There is therefore a more general need to improve the data handling skills of principals and teachers across the board.

Finally, school leadership is the key agent to ensure that school evaluation translates into school improvement. As a result, priority should be given to ensuring school leadership focuses on goal-setting, assessment and evaluation, supports an evaluation culture, and assumes responsibility for instructional leadership (Pont *et al.*, 2008).

Publish externally validated school evaluation reports to complement the publication of national test data

The nature of transparency is a vital issue for policy. Access to credible information about school performance has been a growing phenomenon in recent years. The No Child Left Behind policy in the United States uses both testing and transparency as key drivers of improvement. Inspection reports in the UK countries, Sweden and other European countries are published, including in some cases not just by making them available on a website but by actively sending reports to parents, politicians and the media. The current pattern in Australia varies across jurisdictions although the *My School* website does provide a very public evaluation of school performance in literacy and numeracy.

School reports are made available but the extent to which they contain explicit and independently verified evaluations of the school's performance is very limited. Given the publication of comparative national test data, there remains a strong case to provide complementary evaluative information which broadens the base of evidence and provides more explanation of the factors which have influenced performance. Arguably, testing can only provide a *post hoc* evaluation of performance but good school evaluation is more proactive and should help to identify those factors which are influencing performance at an earlier stage. Consideration should therefore be given to not only continuing to refine and extend the content of the *My School* website but to include direct links to school reports which are validated by external involvement, are more comprehensive in their scope, look inside the "black box" of the working of the school and set a clear and specific improvement agenda.

The role of the media in using the results of evaluations, both quantitative and qualitative, remains problematic. Perhaps the greatest fear of schools in relation to evaluation evidence is the creation of what they regard as simplistic league tables which

rank schools in ways which fail to take account of the factors which influence performance – these have been published in the media for some years (and prior to *My School*) through freedom of information requests. Reactions to the publication of NAPLAN results on the *My School* website at the school level are an example of such concerns. A consistent theme in such complaints is the "crude" nature of the data and the failure to provide the kind of contextual background which is needed for a sophisticated interpretation (although *My School* provides nationally consistent information not previously available as a basis for the league tables). Proposed improvements to the website will address some of these concerns but easy access to straightforward school evaluation reports which provide more of a narrative would take the approach further. There will, inevitably, remain those who are opposed in principle to information about performance being made public at all but much of the current scepticism could be addressed by adopting a more rounded approach to evaluation.

Notes

1. According to TALIS, only 6.8% of Australian teachers of lower secondary education worked in schools where no school self-evaluation was conducted in the previous five years (6th lowest figure among 23 countries against a TALIS average of 20.2%). The corresponding figure concerning external school evaluation is 21.2% (10th lowest figure among 23 countries against a TALIS average of 30.4%).
2. According to TALIS, 86.9% of Australian teachers of lower secondary education are in schools whose principal reported that student test scores were considered with high or moderate importance in school self-evaluations or external evaluations (5th highest figure among TALIS countries, against an average of 76.2%). The corresponding figures for the "use of retention and pass rates of students" and "other student learning outcomes" are 81.9% (8th highest figure) and 94.8% (highest figure) respectively.
3. The Review Team did not find documented research evidence on these potential negative effects of the high stakes of NAPLAN results.

References

Australian Government (2008), *Quality Education: The Case for an Education Revolution in our Schools*, Australian Government, Canberra, *www.deewr.gov.au/Schooling/Resources/Documents/Publications/QualityEducationEducationRevolutionWEB.pdf*.

Australian Government (2010), *Country Background Report for Australia*, prepared for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, Department of Education, Employment and Workplace Relations, Canberra, available from *www.oecd.org/edu/evaluationpolicy*.

Department of Education of Western Australia (2008), *The School Improvement and Accountability Framework*, Perth.

HMIE (Her Majesty's Inspectorate of Education Scotland) (2010), *How Good is Our School*, *www.hmie.gov.uk*.

MCEECDYA (2009), *Principles and Protocols for Reporting on Schooling in Australia*, Ministerial Council for Education, Early Childhood Development and Youth Affairs, *www.mceecdya.edu.au/mceecdya/nap_principles__protocols_for_rep_on_school_2009,27896.html*.

Nevo, D. (2002), "School-based Evaluation: An International Perspective", *Advances in Program Evaluation*, Vol. 8, Elsevier Science Ltd, Oxford, United Kingdom.

Pont, B., D. Nusche and H. Moorman (2008), *Improving School Leadership, Volume 1: Policy and Practice*, OECD, Paris.

Rosenkvist, M. (2010), "Using Student Test Results for Accountability and Improvement: A Literature Review", *OECD Education Working Papers*, No. 54, OECD, Paris, available from *www.oecd.org/edu/workingpapers*.

Chapter 6

Education system evaluation

Monitoring progress towards educational goals is a priority both at the national and systemic levels. This is accomplished namely through the National Assessment Program and state- and territory-based assessments. The monitoring system also includes a Measurement Framework for National Key Performance Measures as well as data and surveys at the systemic level. The strategy draws considerably on public reporting of the progress and performance of Australian students and schools through instruments such as the My School website, the National Report on Schooling in Australia, COAG Reform Council Reports, Report on Government Services in addition to system-level analyses organised through independent reviews. System evaluation builds on a considerable number of strengths: there are clear standard frameworks both for reporting key performance measures and for general government sector reporting; the comparability and coverage of national data are continuously improving; there are strong procedures for system monitoring at the state and territory level; there is transparency in reporting results of national monitoring; and there is extensive use of results from the national monitoring system. Priorities for future policy development include continuing and prioritising efforts to meet information needs for national monitoring; clarifying the role of the National Assessment Program in relation to the Australian Curriculum; further exploiting results from jurisdiction and national monitoring systems for systemic school improvement; and supporting and promoting greater monitoring in the non-government sector.

This chapter looks at system evaluation within the Australian overall evaluation and assessment framework. It refers to approaches to monitor and evaluate the performance of education at the national and systemic (state, territory or non-government system) levels. The main aims of system evaluation are to provide accountability information to the public and to improve educational processes and outcomes.

Context and features

Monitoring national education is a priority for the Australian Government and includes public reporting of the progress and performance of Australian schooling at the core. The rationale is to allow the public to evaluate the system and the Government's performance. Such commitment to transparency has seen significant developments at the national level over a relatively short period and increased collaboration among the states and territories and government and non-government sectors. Demands for comparable information to monitor education outcomes at the national level have increased and new national monitoring tools and authorities have been established.

Increased demand for monitoring education outcomes in Australia

Australia was one of the forerunners in participating in international surveys of student achievement, which provide benchmarking measures of how students in Australia compare on key educational outcome measures with students in other countries.[1] However, it was only in 1989 that ministers agreed on the first set of common national goals for education in Australia (the Hobart Declaration) and committed to an annual report on schooling in Australia from 1990 reporting on "school curriculum, participation and retention rates, student achievements and the application of financial resources in school".[2] This shift to thinking of educational goals at the national level played a major role in establishing a demand for monitoring education outcomes in Australia. The agreed goals listed ten aims including providing an excellent education for all young people, helping them develop self-confidence, self-esteem and respect for Australian and Aboriginal cultural heritage and preparing them to become active and informed citizens, as well as promoting equality of education opportunities and meeting emerging economic and social needs in Australia. Key skills and knowledge were listed and included literacy and numeracy. The national goals were revised in 1999 (the Adelaide Declaration) and new elements included participation in vocational learning programmes, the improvement of learning outcomes for educationally disadvantaged and Aboriginal and Torres Strait Islander students, and access to high-quality education to enable the completion of Year 12 or vocational equivalent. In 1999 a national taskforce was established to develop performance measures to monitor progress toward the national goals. National benchmarks for literacy and numeracy were first reported in the 1999 and 2000 editions (respectively) of the *National Report on Schooling* (MCEETYA, 2000). This corresponded with the shift in focus in general government reporting from inputs to results and the first Budget report on an accrual-based outcomes and outputs framework in 1999-2000 (Australian National Audit Office/CPA Australia, 2008).

The current national goals date from 2008 and are stated in the Melbourne Declaration on Educational Goals for Young Australians. The new national goals continue to place young Australians at the centre with the aim that they become "successful learners, confident and creative individuals, and active and informed citizens" and that schooling should promote both equity and excellence (Curriculum Corporation,

2008) (see Chapter 1). Regarding "equity", there is explicit mention to ensure that: "the learning outcomes of Indigenous students improve to match those of other students" and "socio-economic disadvantage ceases to be a significant determinant of educational outcomes".

At the highest political level there has been recognition of the importance of securing high-quality educational opportunities and outcomes for Australian students. As in many other OECD countries, politicians cite the importance of education's role in securing the nation's future productivity and international competitiveness. The Council of Australian Governments (COAG) estimates that, in combination with early childhood and skills and workforce development policies, improved education policies could boost productivity by up to 1.2% by 2030. As such, education has a prominent place in the 2008 COAG National Productivity Agenda for reform. COAG has set three major targets for schooling including an increased proportion of young Australians attaining senior secondary education and two targets to reduce the performance gap of Indigenous students (see Table 1.2, Chapter 1).

Such high and admirable ambitions for Australian schools have placed increased demand on the national monitoring of education outcomes and national progress toward related goals and targets.

Major stakeholders in monitoring education outcomes in Australia

Responsibility for monitoring and reporting

MCEETYA has held the major responsibility for reporting on national education producing, for example, the 2000 to 2008 editions of the *National Report on Schooling in Australia*.[3] Following its creation ACARA took over responsibility for producing the *National Report on Schooling in Australia* in 2009. Further, in 2010, ACARA assumed responsibility for the National Assessment Program (see below). Thus, ACARA brings together the major functions of monitoring national educational outcomes.

MCEECDYA monitors the work of ACARA on developing the Australian Curriculum and reporting results from the national monitoring system. MCEECDYA's Aboriginal and Torres Strait Islander Education Working Group leads the monitoring and reporting on the Aboriginal and Torres Strait Islander Education Action Plan 2010-14 supporting all education systems to work together towards achieving the targets to reduce Indigenous student achievement disadvantage (Australian Government, 2010a).

In the Melbourne Declaration the Australian Government and the state and territory governments committed to "working with all school sectors to ensure that public reporting: focuses on improving performance and student outcomes; is both locally and nationally relevant; is timely, consistent and comparable" (Curriculum Corporation, 2008). Although the major responsibility for monitoring educational outcomes lies at the school system level, the National Education Agreement (NEA) also clarifies that state and territory governments have responsibility for monitoring all schools, specifically with responsibility for: "the regulatory framework for all schools, including registration and accreditation, educational quality and their performance in educational outcomes, in monitoring and reviewing performance of school systems". The COAG Reform Council plays a key role in monitoring progress of states and territories in meeting targets set in the NEA and publishes an annual performance report on this (see for example COAG Reform Council, 2010).

The Australian Government Productivity Commission is an independent statutory authority with a major role in monitoring education outcomes in Australia and acts as the Secretariat for the Steering Committee for the Review of Government Service Provision. This authority enjoys a fair degree of freedom, although the Government commissions enquiries on a range of economic, social and environmental issues. All Commission members are statutory appointees. Its role is "to help governments make better policies, in the long-term interest of the Australian community" (Productivity Commission, 2009). The Productivity Commission monitors national education and other government sectors on a set of agreed indicators in the annual *Report on Government Services*.

Responsibility for compiling key information for national monitoring

The Australian Bureau for Statistics (ABS) provides much of the data used in system-level reporting. There is an annual collection of nationally comparable data on student and staff in primary and secondary schools by the National Centre for Education and Training Statistics (NCETS) within the ABS. Data for this National Schools Statistics Collection (NSSC) are provided by state Education authorities for government schools and by the Australian Government (DEEWR) for non-government schools according to agreed standard definitions provided by ABS.

Both the Australian Council for Educational Research (ACER) and the National Centre for Vocational Education Research (NCVER) provide key input to the national monitoring systems, as contractors of the Australian Government and state and territory governments. ACER produces a national report on the major results for Australia in the Programme for International Student Assessment (PISA), as well as various related analytical reports, plus provided analysis of variations among schools and states and territories on NAPLAN 2009 for the COAG Reform Council's annual report. NCVER compiles statistics for MCEECDYA on vocational education and training (VET) in schools (see, for example, NCVER, 2009).

National monitoring system

The Australian National Assessment Program (NAP) includes a suite of national and international tests used to monitor progress towards the national goals for education.

Australia has a well-established tradition of participating in international assessment tests and currently participates in two major cyclical surveys: the International Association for the Study of Educational Achievement (IEA)'s Trends in International Mathematics and Science Study (TIMSS) which assesses students in Years 4 and 8 every four years; and the OECD's Programme for International Student Assessment (PISA) which assesses 15-year-old students in reading, mathematics and science literacy every three years. In addition, Australia participated for the first time in the 2011 IEA's Progress in International Reading Literacy Study (PIRLS) which assesses students in Year 4. Participation in such assessments provides benchmark information for states and territories nationally and internationally and allows assessment of progress towards the COAG key outcome that Australian students excel by international standards. A major advantage is that such results allow a monitoring of progress across time, for example, trend data are available for TIMSS from 1995 and for PISA from 2000.

The suite of national assessments includes cyclical sample surveys to monitor student outcomes in science, ICT, civics and citizenship. These tests draw on a statistically

representative sample of students at target year levels (equivalent to about 5% of the corresponding population). Each area is an agreed national priority and is tested once every three years. The first survey was run in 2003 for science, in 2004 for civics and citizenship, and in 2005 for ICT. Each assessment results in a national report showing student average performance and proportion of students at the set "proficient standard" for each state and territory, each school sector and for selected student subgroups (*e.g.* Indigenous, socio-economic background) and allows a reporting of progress over time, as each subject is assessed every three years (see for example MCEECDYA, 2010). For both ICT and civics and citizenship students are assessed in Years 6 and 10. Scientific literacy is not assessed for Year 10, but given that the "PISA definition formed the basis of the work to assess the scientific literacy of Year 6 students", results from the PISA Science assessment serve as a measure later in schooling (Curriculum Corporation, 2004).

The National Assessment Program – Literacy and Numeracy (NAPLAN) comprises annual full-cohort tests in reading, writing, language conventions (spelling, grammar and punctuation) and numeracy for students in Years 3, 5, 7 and 9. NAPLAN tests are designed to measure student progress and accordingly use one common scale of performance bands (for more information, see Chapter 3). In this way, it is possible to gauge student progress in the national tests on a subsequent year, for example, it will be possible to see how well a student performs on the common NAPLAN reading scale at four different stages of his or her schooling (in Years 3, 5, 7 and 9) – such progress reporting will commence in 2010. The 2008 and 2009 results already allow comparison of state and territory performance against national average and were extensively reported in the main national monitoring reports. Results are reported against national minimum standards defined for each of the five areas. Reports summarise performance (and confidence intervals) relative to students' mean scale scores and percentages scoring at, above, and below national minimum standards for each year level. Performance breakdowns are provided by gender, Indigenous status, language background other than English (LBOTE), geographical location (metropolitan, provincial, remote and very remote), geographical location by Indigenous status, parental education and parental occupation.

Strengths

Common reporting frameworks well established

A core strength of the evaluation of education in Australia is the existence of clear standard frameworks both for reporting key performance measures and for general government sector reporting.

MCEECDYA since 2000 has worked on producing the *Measurement Framework for National Key Performance Measures*. Ministers first defined national Key Performance Measures in early 2000 as "a set of measures limited in number and strategic in orientation, which provides nationally comparable data on aspects of performance critical to monitoring progress against the National Goals for Schooling in the 21st Century" (MCEETYA, 2008). This framework clearly presents the agreed measures and their source for each of the priority areas: literacy, numeracy, science literacy, civics and citizenship, information and communication technologies (ICT) literacy, vocational education and training (VET) in schools, student participation, student attainment and student attendance. The core of the framework is a schedule setting out key performance measures and an agreed assessment and reporting cycle for the period 2006-2014. In

2008, the framework was enhanced by the inclusion of comparable measures on literacy and numeracy from NAPLAN. The framework was reviewed and in late 2010 was further refined to incorporate the full suite of agreed national key performance measures, including the COAG measures.

The *Report on Government Services' Performance Indicator Framework* provides a common reporting basis for each government sector. A recent independent review of the framework highlighted the potential efficiency of sharing information on performance indicator methodology as well as some performance measures across government services (Steering Committee for Review of Government Service Provision, 2010).

Strengthened set of national monitoring tools

National information to monitor education in Australia is largely compiled by the Australian Bureau of Statistics (ABS) in collaboration with the different states and territories and the Australian Government. Along with ongoing work to improve the quality of national statistical indicators, the addition of comparable outcomes information from the National Assessment Program – Literacy and Numeracy (NAPLAN) has significantly strengthened the set of national monitoring tools.

Improving comparability and coverage of national data

The ABS National Schools Statistics Collection (NSSC) has been refined over the years: The first such collection was conducted in 1981, but only covered government schools at that time – data for non-government schools were collected as of 1984. In 1989 data were collected to distinguish Aboriginal and Torres Strait Islander pupils. Most recently, efforts have been made to address concerns on one of the key indicators "Apparent Retention Rates" and two new indicators were published "Full-time plus part-time School Participation Rates (SPR)" and "Apparent Progression Rates (APR)" in *Schools, Australia, 2009*. This follows a proposal made in an ABS research paper to make better use of existing data collected via the NSSC (Australian Bureau of Statistics, 2006).

Nationally comparable data on student outcomes

Since 2001, results from the international assessment PISA have been effectively reported at the national level, providing comparable measures on student outcomes in reading, mathematics and science at age 15. A national report is published which includes results by school sector (from PISA 2009) and by state and territory (see, for example, Thomson *et al.*, 2011, and Lokan *et al.*, 2002). Such reporting allows a benchmarking of states and territories both nationally and internationally.

Up until 2007, national reports included information on student performance in literacy and numeracy that was drawn from annual state and territory assessments, these had a fairly good coverage of students in the government and non-government sectors (see Table B1, MCEETYA, 2007). To provide comparability, results were equated through a national process and a national benchmark was established. However, in 2008 common national tests in core skills (NAPLAN) were conducted for the first time for all students in all school sectors and replaced previous state and territory assessments. In this way, for monitoring purposes, NAPLAN provided a more robust measure of skills across states and territories and did not impose additional testing requirements on schools. According to the ACARA website, NAPLAN "has provided consistency, comparability

and transferability of information on students' literacy and numeracy performance nationally".

From 2010 onward, NAPLAN will provide measures of student gains which can be aggregated to national and state and territory levels. ACARA reports that a rigorous equating process is undertaken each year to ensure that NAPLAN results can be compared from year to year. This allows monitoring of schools and systems over time.

A strong and stable set of national measures offers the advantage of being able to "weather" changes in political systems at the state and territory level. State and territory measures may become more aligned with national measures and thus more stable. For example, Victoria has revised its early years English diagnostic tools (the English Online Interview) to align with both Victorian Essential Learning Standards (VELS) and the Year 3 NAPLAN literacy scale.

Strong procedures for system monitoring at the state and territory level

Strengthening structures and capacity to monitor system performance

Several jurisdictions have strengthened structures to monitor schooling over recent years. For example, new structures have been introduced to monitor the government school sector. South Australia established twelve regions with Regional Directors and support teams to provide "supportive, enabling leadership" to schools. In 2009 new consultancy positions were created in performance analysis and reporting to ensure regions develop a focus on using data to inform improvements (Department of Education and Children's Services, 2010). Similarly in 2010, the Australian Capital Territory established a new School Improvement Division and four regions each with a School Network Leader to promote accountability and networking (ACT, 2010) and Queensland created seven regions which include regional leaders providing a single point of accountability and established new performance frameworks (Department of Education and Training, 2010). Victoria aims to "pursue a stronger systemic approach to school improvement in government schools based upon driving improvement through the role of regional networks, and stronger interventions in schools where performance needs to improve" (State of Victoria, 2008).

As part of the Smarter Schools – Literacy and Numeracy National Partnership, Catholic schools in the Kimberley region of Western Australia have introduced support structures by training regional consultants and key school staff in data analysis to better understand student literacy and numeracy skills (Australian Government, 2010a).

Use of ICT for reporting systems

Most jurisdictions offer software to aid schools in fulfilling their reporting responsibilities. For example, in New South Wales, 500 government schools were offered software to produce student reports in 2009. "School Based Student Reporting Version 4" provides online access to central enrolment and registration systems (New South Wales Department of Education and Training, 2010). Similarly, in Western Australia, the "School Information Management System" is used by all but five government schools and includes information on finances, aspects of teaching and learning and other reporting.

Some jurisdictions also report summaries of school performance that are made publicly available on line. For example "Schools Online" in Western Australia provides information on the schools including performance indicators on attendance, literacy and numeracy and for senior secondary schools, senior secondary qualifications and Year 12 destination. This is an interactive software allowing users to navigate to particular schools and download different information about the school. In Victoria, school performance indicators are available on line as static performance sheets (pdf files) with absolute results for the given school compared to other government schools in Victoria, and also with a simple indicator (lower, similar, higher) of how these results compare to similar schools (with a similar academic intake, socio-economic composition, number of Indigenous, non-English speaking and refugee students, number of students with a disability and the size and location of the school) (see also Chapter 5).

State monitoring tools

While the introduction of NAPLAN replaced the previous eight literacy and numeracy tests in states and territories, there are examples of complementary monitoring tools to shed more light on specific needs. In Western Australia, as NAPLAN replaced the Western Australian Literacy and Numeracy Assessment (WALNA), resources were given to establish two new assessments (the Western Australian Monitoring Standards in Education) in different areas: science, and society and environment (Years 5, 7 and 9). These are standards-based tests that allow the tracking of changes in performance at the system, school and student level. In South Australia as part of a focus to improve mathematics outcomes for Indigenous and socially disadvantaged students, Years 3 to 5 students in selected schools sat a Progressive Achievement Test in Mathematics at the start and end of the academic year. Results indicated some progress in reducing achievement gaps for these student groups (Department of Education and Children's Services, 2010). South Australia also planned to release in 2010 a "Student attendance and behaviour management data warehouse" which will allow analysis of student absenteeism and behaviour against their literacy and numeracy performance (Department of Education and Children's Services, 2010). Queensland introduced Comparable Assessment Tasks in science, mathematics and English in Years 4, 6 and 9 (Department of Education and Training, 2010).

Most jurisdictions also systematically collect qualitative feedback from the primary users of education, that is, the students and parents. An example of effective collaboration is that Tasmania introduced the annual student and parent opinion surveys used in Victoria from 2007 onward (Department of Education, 2010a). Independent Schools Victoria offers its members a set of surveys for parents, school staff, students and school boards/councils.

There are also attempts to monitor student transitions after the completion of Year 12. In Queensland the "Next Step" survey provides information on Year 12 students' expected destination for government and non-government schools. For example, in 2009, 36600 graduates out of 44500 completed the survey (Department of Education and Training, 2010). In Victoria, the Department of Education and Early Childhood Development runs the "On Track Year 12 Completer Survey" and has given greater responsibility to schools to monitor and evaluate school interventions to students who are at risk of dropping out via the "Student Mapping Tool" (State of Victoria, 2009).

External reviews of school systems

Government audit offices are increasingly moving beyond compliancy reporting to auditing performance management. For example, the Victorian Auditor General's Office has conducted audits on how the Department of Education and Early Childhood Development monitors government school performance, on the Victorian accountability framework, as well as on literacy, numeracy and student well-being. In 2010 the Queensland Audit Office conducted a review of the Department's systems to use student information to inform literacy and numeracy teaching and learning (see Queensland Audit Office, 2010). Further, Queensland has commissioned an ongoing external evaluation of the Queensland Curriculum, Assessment and Reporting Framework. The Catholic Education Office in Sydney commissioned an external review in 2004 on how appropriate and effective their services were in improving education standards in schools within the system (Australian Government, 2010b).

Transparency in reporting results of national monitoring

All reports on education in Australia are publicly available on various websites. National and state/territory level statistics have been reported in the MCEECDYA national report series and this continues under ACARA's management. Reports up until 2008 can be found on the MCEECDYA website where there is also much information on helpful background to the monitoring system, for example successive reports on the measurement framework and technical reports advising on the development of the national measures. Similarly, the *Report on Government Services* series produced by the Productivity Commission is available on their website. During the OECD Review, discussions with stakeholders indicated that these reports have a fairly high degree of visibility and are deemed to be of high quality and present relevant evidence/facts. Due attention is paid to improving alignment in reporting at the national level. For example, the *2010 Report on Government Services* already reported a revised set of indicators on education and training in alignment with the National Education Agreement and the National Agreement for Skills and Workforce Development. Further, NAPLAN results for each school are published on the *My School* website (see Chapter 5). As a commitment to timely publication of results, ACARA reports an initial overview report with major results on each of the five NAPLAN scales for national and state and territory levels (usually published in September) and a more in-depth report of results by different student groups is published later (usually December).

Extensive use of results from the national monitoring system

Many Australian government bodies make use of the results from the national monitoring system. The annual report from DEEWR includes performance indicators that draw on results from NAPLAN and the national monitoring surveys, as well as enrolment and apparent retention rates. Results are also extensively reported in the *Report on Government Services* and the *National Report on Schooling in Australia*. Indicators in these reports are also extensively reported as part of state and territory government reporting (see below).

Use of results from international assessments

According to stakeholders during the OECD Review, PISA results (both national and for states and territories) have garnered significant policy attention and served to motivate

educational policies to improve student performance. These have also drawn attention to the average performance disadvantage for students from less advantaged socio-economic background and fed into the national goals and COAG agenda. In their annual performance reports, Queensland uses results from PISA and TIMSS to judge progress in science performance and Tasmania uses PISA data to judge the effect of socio-economic background on literacy and numeracy outcomes (Department of Education and Training, 2010; Department of Education, 2010b). Also, international assessments have informed the debate about the new Australian Curriculum. For instance, ACARA considered the curriculum of other countries, including those that perform highly in international assessments, such as Finland, Canada, Hong Kong and Singapore.

Using national monitoring results to report on school system performance

All government departments produce an annual report on major activities, including both financial and performance information. In performance reporting, the major focus is on performance outcomes for the government school sector, although the reports also usually include minimal reporting on the non-government school sectors (*e.g.* enrolment figures, new schools registered, proportion of schools meeting agreed requirements). A common feature in the 2009/10 government reports is the prominence of NAPLAN data in the performance monitoring. The exact format for reporting NAPLAN results varies according to the emphasis on different monitoring goals in each state and territory. The majority of jurisdictions report according to the national minimum standard. In New South Wales and Queensland this is the proportion of students at or below the minimum standard and in the Northern Territory, South Australia, Tasmania, Victoria and Western Australia, this is the proportion of students achieving the national minimum standard – as reported in the *Report on Government Services*. Tasmania reports the full distribution of student performance in each of the NAPLAN bands. The Australian Capital Territory, however, simply reports the average performance of students. New South Wales is the only government department to focus on the proportion of students performing in the top two bands of NAPLAN. Also, New South Wales reports results from the national assessment in civics and citizenship.

For the Catholic sector, there are also annual reports produced in some jurisdictions. For example, the Catholic Education Commission of Victoria included NAPLAN results on the percentage of students who had achieved results at or above the national minimum standard in each year (Catholic Education Commission of Victoria Ltd, 2010).

Systemic use of national monitoring results for school improvement

Stakeholders in Queensland informed the OECD Review Team that relatively low performance on NAPLAN had stimulated reform processes, and representatives of Indigenous populations and advocacy groups noted the role of NAPLAN in highlighting performance gaps and motivating new programmes. In Victoria, there are examples of use of NAPLAN results to monitor both the government and Catholic systems. The Department of Education and Early Childhood Development makes use of NAPLAN results to monitor the consistency of teacher grading across schools. Where statistical checks reveal discrepancies, this is followed up in school reviews. The Catholic School Commission analysed NAPLAN results and determined that teachers were underestimating student performance. Such analysis has driven forward promotion of differentiated teaching and working with student performance data by teachers in Catholic schools. The Catholic sector reported to the OECD Review Team that it has

invested heavily in developing teacher capacity to work with student performance data. NAPLAN results also provide a primary mechanism for establishing targets and incentives for the accountability of schools which are part of the National Partnership on Literacy and Numeracy (*e.g.* results linked to reward payments) and figure prominently in states and territories plans for school review and improvement.

Promoting use of NAPLAN results throughout the system

There is a high level of feedback to schools on their student performance in NAPLAN to help promote use of such results to improve student outcomes within schools. Victoria provides a data service for schools offering comprehensive information on NAPLAN results for different groups and by each different question on the test. For example, schools can group results by class or gender and analyse how different students performed on different parts of the tests. The School Measurement Assessment and Reporting Toolkit (SMART) is used in New South Wales, the Australian Capital Territory and South Australia and allows extensive analysis of performance on the NAPLAN tests by student, groups of students, class and school. SMART was developed by the New South Wales Department of Education and Training and offers teachers a sophisticated tool to analyse their student performance and understanding of key areas covered in the NAPLAN tests, plus is a useful resource of teaching strategies and related worksheets for teachers and sometimes students (see Box 2.2 in Chapter 2). In Western Australia, the Student Achievement Information System is an analytical tool for teachers to track and graph individual and group student achievement data over time and can also be used at the school and system level to moderate grades and review courses (Department of Education, 2010b).

Challenges

There are some gaps in the national monitoring system

A core ambition in the national goals for schooling is to keep them relevant in the context of emerging economic and societal demands. By definition, this poses a challenge to a monitoring system to keep track of emerging priority demands. While the significant progress that has been made in strengthening monitoring at the national level (in particular the sheer speed of this achievement and level of collaboration with stakeholders) is commendable, there remain challenges in some key measurement areas. National measures are not available for all national goals and, in particular, data for some of the key student subgroups (*e.g.* Indigenous and socio-economically disadvantaged students) suffer in terms of coverage and quality at the national level. Strong support at the highest political level has driven much of the progress in the strengthened monitoring system. However, this has also brought challenges. The Productivity Commission attributes some of these measurement challenges to the top-down approach of the COAG agreement which gave rise to issues such as how to report against some of the criteria, for example, student engagement and early childhood education outcomes. This was echoed by representatives of the Australian Education Union who expressed concern that the national approach had not been adequately informed by input from educational specialists.

The COAG Reform Council (2010) asserts that a “lack of comparable, timely data limits the council’s ability to undertake its role of performance monitoring”. For example,

the Council identifies weaknesses in the indicators to monitor the first of five COAG key outcomes "All children are engaged in and benefiting from schooling". First, attendance data are not adequate for comparison of states and territories or developments over time: "the data cannot be added or averaged to provide a national figure and cannot be compared across jurisdictions or sectors". Second, the current indicator does not monitor the extent to which students are "benefiting from schooling".

A national measure on individual student or school socio-economic background is incomplete

Given the prominence of equity in the national performance framework, concerns about the quality of the nationally available measures on student socio-economic background present a significant challenge. This is something that MCEECDYA and national stakeholders are well aware of. A report commissioned by MCEETYA in 2000 argues strongly for the improvement of national data in this area and warns against the use of an area-based measure of socio-economic status (Marks *et al.*, 2000). The report argues that the use of an area-based measure of socio-economic status to estimate an individual's socio-economic background: is subject to considerable misclassification error, especially in regional and rural areas; is not cost effective; often relies on out-of-date information; undermines conclusions about between-system and over-time differences in the importance of socio-economic background on educational outcomes; cannot be used to categorise individual socio-economically disadvantaged students when reporting student outcomes; and does not allow analysis "controlling" for differences between different student groups, *e.g.* Indigenous and non-Indigenous students. Currently the National School Statistics Collection does not collect any information related to socio-economic status at either student or school levels (COAG Reform Council, 2010).

The quality of some of the completion data is of concern

There is a lack of data on participation and completion at the national level. Currently, such information is taken from ABS survey data and in some cases does not allow finer breakdowns by state and territory, government and non-government sectors or for Indigenous students. For example, a key performance measure taken from the ABS Survey of Education and Work is the "percentage of 20-24 year-olds with Year 12 certificate or equivalent vocational qualification". Such a measure does not allow adequate monitoring of changes across time for smaller jurisdictions. The Australian Capital Territory Department of Education and Training reports that although data for 2009 were 80% and for 2008 were 88%, due to the small sample size such change cannot be said to be significant (ACT, 2010). Of note also, results for the Northern Territory and Tasmania have standard errors from 25 to 50%. These data limitations are signalled by the COAG Reform Council (2010).

Currently, administrative data on Year 12 completion are not comparable across states and territories.[4] However, administrative data would be timelier and could be more complete than the ABS survey data and as such the COAG Reform Council proposes to include an additional attainment indicator based on the administrative data. This would require work on the comparability of state and territory administrative data (COAG Reform Council, 2010).

Further steps are needed in monitoring the new Australian Curriculum

The current suite of national assessments predates the Australian Curriculum. The first phase of implementing the Australian Curriculum is almost complete and covers the learning areas of English, mathematics, science and history. Currently, an assessment of history is not included in the National Assessment Program. Similarly, there is ongoing work to introduce an Australian Curriculum for languages, geography and the arts, which are also not covered in the National Assessment Program. While some areas of English, mathematics and science are nationally assessed, the assessments are not aligned with the Australian Curriculum. The curriculum in each of these cases places emphasis on standards. However, it is not clear to what extent the current suite of assessments will reflect performance by these standards. Further, the current assessment of literacy and numeracy covers a relatively narrow content area, as each child sits exactly the same test (see also Chapter 3). In this context, it should be noted that a review of the National Assessment Program has been mandated and will occur once the Australian Curriculum is implemented.

There is room to improve the use of results from the national monitoring system

The abundance of new information from the national monitoring system offers many opportunities to engage stakeholders in supporting student outcome improvements. However, this has increased demands on reporting of such information. The Australian Parents Council survey in 2008 (although only with a 30% response rate) showed that parents asserted the right to a wide range of information about schools and about school and student performance. However, while parents demand to see performance comparisons between schools and students, their understanding of the concepts behind the design of such comparisons is incomplete. National representatives from industry groups reported to the OECD Review Team that the increasing complexity of some outcome reporting had not been accompanied by good guidelines for parents and employers. The technical aspects of assessment are often not understood by these important stakeholders. Furthermore, the Year 12 results are described differently across the states and territories. The New South Wales Department of Education in its 2009 annual report notes that it plans to establish a Standards and Assessment Framework Working Group to review and revise information in the NAPLAN student reports to make them clearer and easier for parents to understand (New South Wales Department of Education and Training, 2010).

In particular, there seems to be potential to further exploit results from the national monitoring sample surveys in science, ICT and civics and citizenship. Most Department Annual Reports do not include this information in their performance monitoring. In general, based on stakeholder meetings during the OECD Review, these assessments are not highly visible, although there were concerns expressed about schools' attention to ICT literacy, presumably based in part on prior national, state and territory results in this area.

There are varied practices among states and territories in monitoring schools across different sectors and in systems of data collection

The Melbourne Declaration places strong emphasis on the fact that Australian governments "commit to working with all school sectors" on all the key areas for schooling. Notably, this includes timely, consistent and comparable public reporting (Curriculum Corporation, 2008). Due to differing systems of data collection and degree of analytical sophistication among different education systems, there is a two-fold challenge: to engage a high-quality common core of monitoring indicators and to allow continued

freedom for diversity in systems of indicators. At the heart of this is the challenge to use the best available evidence at the local level to more readily assess and evaluate innovation and not to allow sometimes "richer" information systems to be overshadowed by the core national indicators. Rather, the challenge is to ensure that the national monitoring system benefits from the existence of more sophisticated indicators in some education systems. One example of where data in the national system are limited is information on student socio-economic background that is based on their postcode (*i.e.* area-based), whereas better quality information on individual student socio-economic background may be available in some jurisdictions. An additional challenge is that even where there are better quality data available, these may only have partial coverage. For example, the Victorian Department for Education reported that there is socio-economic background information based on parental education and occupational status available for well over 90% of children in government and Catholic schools, however, this is not the case for Independent schools. The variation in coverage and comparability of administrative data collection poses significant challenges to the national monitoring system (see above).

While the National Education Agreement clarifies that the states and territories have responsibility to monitor all schools, the extent to which corresponding agreements and systematic collection of information from the non-government sectors are in place varies considerably. The monitoring of non-government sectors is generally conducted via state or territory regulatory authorities, but reporting on their outcomes is still limited to a simple set of compliance statements and does not focus on performance. Schools from all sectors are required to publish an annual report, including information on their performance. However, it is unclear to what extent such information for non-government schools is aggregated to the system level and analysed.

Policy recommendations

The OECD Review Team commends the current system to monitor key educational outcomes. Based on clear reporting frameworks linked to agreed national goals, the monitoring system has been significantly strengthened by the addition of nationally comparable measures on student numeracy and literacy outcomes. In this context, the OECD Review Team outlines four policy recommendations to further strengthen system evaluation in Australia:

- Continue and prioritise efforts to meet information needs for national monitoring;
- Clarify the role of the National Assessment Program in relation to the Australian Curriculum;
- Further exploit results from jurisdiction and national monitoring systems for systemic school improvement; and
- Support and promote greater monitoring in the non-government sector.

Continue and prioritise efforts to meet information needs for national monitoring

The OECD Review Team endorses the two priority areas identified by the COAG Reform Council (2010) to improve performance reporting: "achievement of Indigenous students and students from low socio-economic backgrounds" and "reporting of change over time". The immediate priority for meeting information needs to adequately monitor progress towards national goals is to strengthen the information systems regarding student

socio-economic and Indigenous status. The quality of socio-economic background data, in particular, proves inadequate for monitoring progress on several key indicators. The pilot in two jurisdictions for collecting enrolment, attendance and progression data at the unit record level as part of the National School Statistics Collection should prove helpful in determining whether to implement this throughout administrative data collections.

Several factors point towards the importance of strengthening administrative data collections:

- The inclusion of information on student socio-economic background and Indigenous status would be critical. Case in point, the recent attempts by ABS to report a better quality indicator on apparent retention rates is hampered by the lack of administrative information on either of these elements. The compromise, therefore, is to continue reporting the original indicator based on survey data to allow reporting by Indigenous status. Note that apparent retention rates are heavily reported in state and territory annual performance reports.
- Improved enrolment and registration data will also benefit the very prominent reporting of NAPLAN results – including improving the reporting of results for "like schools" in the *My School* website. The collection of information directly from students on their socio-economic background during the test administration proves challenging: from the experience of student self-reports in NAPLAN, data reported on Indigenous background are reliable, but those on parental education are not (due to large amounts of missing data). On average, for the 2009 NAPLAN tests, 30% of students did not report information on parental education. The NAPLAN 2009 report states that parental education and occupational status information may not have been recorded upon student enrolment.
- More reliable and timely reporting on Year 12 completion via administrative collections would allow comparison across time for smaller jurisdictions and for key student groups. In 2009/10, all states and territories included indicators on Year 12 completion as part of their annual performance report. So improving the comparability of reporting here at the national level would be of significant benefit to all systems providing a national benchmark.

The political commitment to introduce a unique student identifier system in Australia in the long term offers a good opportunity to strengthen reporting on student socio-economic background. Both New South Wales (SMART) and Victoria (Ultranet) have recently introduced Web-based systems with unique student identifiers (see Box 2.2). However, background information on students is not included in the Ultranet, only the students' name and maybe photograph.

Given the central importance on the national agenda of closing the performance gap of Indigenous students, the OECD supports the decision to extend the sampling in PISA to allow reporting of results for Indigenous students by state and territory. This will provide comparable information on the performance of Indigenous students relative to non-Indigenous students in all states and territories and will allow absolute benchmarking of Indigenous student performance internationally and to gauge progress on Indigenous student representation among the best performers internationally. Importantly, much contextual information is collected during the administration of PISA from both students and school principals and analysis of such data should help to shed light on common risk and success factors for Indigenous students. For example, in Denmark there was a special administration of PISA by the Rockwool Foundation Research Unit in 2005, surveying

only schools with high proportions of students from a non-Danish background. This "PISA Ethnic" study was conducted in 112 schools and led to much analysis by national researchers on risk factors for underperformance of different student groups and has seen much attention paid to this issue by national and local policy makers and the education community.

Clarify the role of the National Assessment Program in relation to the Australian Curriculum

The OECD Review Team commends the Australian Government on the well-thought-out National Assessment Program. It was conceived to measure the national goals for schooling and provides information on student outcomes in ICT, science, civics and citizenship, and numeracy and literacy and offers a balance of sample surveys and full cohort tests. The results remain valid and useful measures of progress towards the current national goals feed into policy making at many different levels. In particular, NAPLAN aims to shed light on the national goal that "successful learners have the essential skills in literacy and numeracy…" NAPLAN tests the basic skills and judges how students perform against minimum national standards in numeracy and literacy. Given the importance of NAPLAN in national reporting and the high investment by several education sectors and educators to work effectively with results, the OECD Review Team would caution against changing the current format of these tests.

However, the introduction of the Australian Curriculum poses a new challenge to the existing instruments. It is, therefore, important to clarify the role of the current suite of national assessments in relation to the new Australian Curriculum. For example, it is not clear to what extent the Australian Curriculum for English and mathematics will align with NAPLAN and whether there will be a demand to monitor beyond the essential skills and focus more on the assessment of higher-order thinking skills (see Chapter 3). This is in line with national concerns over observed stagnation at the top performance levels in international surveys. The particular tension that arises here originates in the use of NAPLAN results for rewarding national partnership payments. Literature on the use of standardised test results for reward/sanction largely draws attention to the importance of alignment of the test to the curriculum (see Chapter 3). ACARA and NAPLAN's developers, as noted earlier, are aware of these challenges and plans have now been established to review the entire National Assessment Program in relation to its alignment with the Australian Curriculum once the Australian Curriculum is in place (see Chapter 3).

Further, policy makers in conjunction with ACARA may want to consider ways to further assess the implementation of the new Australian Curriculum. One consideration could be to extend the current cycle of sample surveys to cover new areas of the Australian Curriculum. Green and Oates (2009) identify one of the key issues in monitoring education systems as the effects of innovation and change. Indeed, the Netherlands recently introduced a new sample survey to monitor the progress of a current school reform. This comes in addition to a long-established cyclical sample survey monitoring a broad range of disciplines. Similarly, New Zealand monitors a broad range of disciplines on a four-year cycle (see Box 6.1). An interesting aspect to the New Zealand system monitoring is the engagement of professionals to score student work in the annual national monitoring tests. There is an open call each year for applications from teachers to score student work on test questions that require professional judgement, *e.g.* open-ended questions where students develop their answers. Participating teachers

are paid a fee to score student work over a one-week period. This takes place under the direction of the tests administrators and teachers can work individually or in pairs. It is hoped that participation in such a scoring process would benefit teachers and help them to develop their professional judgement.[5]

Further exploit results from jurisdiction and national monitoring systems for systemic school improvement

States and territories should continue efforts to strengthen monitoring structures, in part by further exploiting the analysis of results from local information systems and the national monitoring system and importantly by ensuring adequate monitoring and follow-up on priority areas (*e.g.* underperforming schools) and the impact of departmental interventions.

The importance of the careful monitoring of school system performance is highlighted in the most recent annual report by the Office for Educational Standards in England (Ofsted, 2010). Ofsted lists four elements to a systemic school improvement strategy: "setting the standard, which is done through inspection frameworks and local and national targets; avoiding any school becoming inadequate, which depends on effective monitoring and accountability; quick turnaround of any school that becomes inadequate; and sustaining good and outstanding practice." In cases of rapid and effective turnaround of schools that had been classified by the Inspectorate as having significant quality concerns and needing intervention ("special measures" schools), this was largely due to quick, decisive intervention by local authorities and careful monitoring and follow-up of schools.

For example, the Queensland Department for Education and Training produces regional performance reports which the Queensland Audit Office (QAO) (2010) assessed to "incorporate a broader range of student and school performance data to complement the NAPLAN data". However, the QAO saw room for the Department to monitor that schools had adequate assessment policies in place, it assessed that regions "were not holding schools accountable for implementing the actions endorsed in school operational plans" and that regional planning and reporting requirements were not clear. Although links between school plans and the Department's strategic plan were clear, this was not always the case for regional plans. Further, the QAO saw no monitoring of implementation of school annual plans and noted that some school principals "treat the process as a compliance exercise rather than a key mechanism to identify improvement goals and strategies".

Box 6.1 Sample surveys in the Netherlands and New Zealand

In the Netherlands, a new monitoring survey – the Annual Survey of Educational Levels (JPO) – was introduced in 2008 to specifically monitor progress on the roll out of the Ministry for Education, Culture and Science's quality agenda "Schools for Tomorrow" and monitors student mastery of Dutch language and mathematics at two points in primary education (Years 4 and 8). Results are reported and analysed for four major regional groupings in the Netherlands. Analysis of performance in urban and rural classifications is also possible (CITO, 2009).

This comes in addition to the existing monitoring sample survey that has been administered periodically in different disciplines since 1987 and monitors skills in Dutch and mathematics on a five-year cycle (Periodical Survey of Education [PPON]). Other curriculum areas that are monitored in the PPON include world studies, history, geography, biology, physics/engineering, English, music and physical education (CITO, 2008). The design of the PPON aims to provide robust measures of changes over time covering large amounts of the curriculum. The design of JPO aims to provide more regular and timely feedback on a narrower area corresponding to the national reform agenda in primary education. Both the PPON and JPO monitoring surveys use Item Response Theory and therefore allow reporting of what students can or cannot typically do against defined performance standards.

In New Zealand, the National Education Monitoring Project (NEMP) established in 1993 assesses students in primary education in two different year groups (Years 4 and 8) and follows a set four-year survey cycle. In this way the NEMP is conducted each year, but assesses a different set of disciplines. For example, in Cycle 2 music, technology, reading and speaking are assessed, and in Cycle 4 listening and viewing, health and physical education, and writing are assessed. These disciplines, therefore, will only be tested every four years. This allows monitoring of a broad coverage of the national curriculum. According to the NEMP website, the purpose of monitoring samples of students at successive points in time is to identify and report trends in educational performance, to provide good information for policy makers, curriculum specialists and educators for planning purposes and to inform the public on trends in educational achievement.

Sources: CITO (2008, 2009); *http://nemp.otago.ac.nz.*

Implementing efficient information systems is a first step and many education systems in Australia have made significant investments in this area (*e.g.* SMART, Ultranet) (see Box 2.2 in Chapter 2). The implementation of analytical software for working with NAPLAN results and related training of professionals in how to effectively use this should bring wider benefits of further promoting the use of data by educators. Equally, building capacity at the state, territory and regional levels to work effectively with these results should bring several benefits. A study by the United States Department of Education (2010) highlights the need to design links among information systems to be able to analyse the impact of particular educational programmes or interventions (Box 6.2). Further, it highlights the important role that school districts play in promoting schools to work with data effectively. Box 2.3, in Chapter 2, presents the example of a focused body within the central department in Ontario to promote and build capacity throughout the education system, which draws in part on an information system that allows the monitoring of the impact of particular initiatives introduced by the education department.

Box 6.2 Use of data systems for decision making by educational districts in the United States

A study by the United States Department of Education (2010) that ran during 2006-2008 examined how education data systems varied across educational districts and how they were used to aid decision making. The report uses the Wayman (2005) classification of four types of electronic student data systems:

1. Student information systems providing real-time access to student data on attendance, enrolment, grades and schedules;
2. Data warehouses providing access to current and historical data on students, finances and staffing;
3. Instructional or curriculum management systems providing planning tools, links to state content or performance standards and communication tools; and
4. Assessment systems supporting the organisation and assessment of benchmark data.

In general there was a huge increase in reported availability of data systems. Virtually all school districts had student information systems storing basic information on enrolments and attendance and 79% reported having an assessment system to organise and analyse benchmark assessment data. The least common system was on instructional or curriculum management (64% of school districts). The major challenge reported by school districts was to link these multiple data systems to better support decision making and in particular to better link student data to instructional practice. The report found that most systems had developed in response to accountability requirements and less than half the school districts could link outcomes to processes in order to monitor and promote continuous improvement. An example here is that only 42% of school districts could link student performance to participation in particular programmes. The most common school district policies to promote schools to use data was to incorporate this in school improvement planning, providing professional development activities and support positions for system implementation and developing data generation and analysis tools. Examples of support provided by school districts included: technical expertise to schools, "data coaches" available to schools, creating easy-to-read data "dashboards" to make information more accessible to teachers, and developing benchmark and formative assessments providing teachers with more timely data on student progress.

Source: United States Department of Education (2010).

Support and promote greater monitoring in the non-government sector

The OECD Review Team notes the high degree of collaboration among the government and non-government sectors in many states and territories. There may be ways to more efficiently meet state and territory government responsibilities for "timely, consistent and comparable reporting" in all school sectors. Strengthened administrative data collections would make a key contribution to this end. Another possibility is for states and territories to establish common performance summary reports for schools in all sectors. This is planned in Victoria (see Box 6.3). Another possibility would be to include the monitoring information on non-government sectors as part of the annual government education department reporting.

Box 6.3 Proposed monitoring of schools in all sectors in Victoria

As part of its responsibility to monitor minimum standards, the Victorian Registration and Qualifications Authority (VRQA) is responsible for ensuring that all schools monitor and report on student performance and provide information on student attendance and performance and school finances to the school community (State of Victoria, 2009).

The Victorian Blueprint for Education and Early Childhood Development (State of Victoria, 2008) states that the Department will "promote partnerships between government and non-government schools, consistent accountability frameworks and greater transparency about performance and provision from all schools regardless of sector". A report on the implementation of this (State of Victoria, 2009) indicates that a common reporting performance summary will be prepared for each school using data available in both the government and non-government sectors (NAPLAN, enrolment, Victorian Certificate of Education, Victorian Certificate of Applied Learning and Vocational Education and Training, "On Track" student destination data and the International Baccalaureate).

Sources: State of Victoria (2008, 2009).

Consideration could also be given to extending the mandate of state and territory Auditor General Offices to the review of all schools receiving government funding. For example, the Victorian Auditor General's Office's (VAGO) mandate is currently limited to the government school sector. This is also the case for audit offices in New South Wales and Western Australia.

Notes

1. Australia was one of 11 education systems participating in the 1963-67 First International Mathematics Study (FIMS).
2. The national goals can be found at:
   Hobart Declaration:
   www.mceecdya.edu.au/mceecdya/hobart_declaration,11577.html;
   Adelaide Declaration:
   www.mceecdya.edu.au/mceecdya/adelaide_declaration_1999_text,28298.html;
   Melbourne Declaration:
   www.mceecdya.edu.au/verve/_resources/National_Declaration_on_the_Educational_Goals_for_Young_Australians.pdf.
3. Reports are publicly available from *www.mceecdya.edu.au/mceecdya/anr*. Note that since 2009, MCEETYA was split into two different councils. The council responsible for schooling is now the Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA).
4. In addition, data may be inflated in the case that a student receives multiple qualifications, *i.e.* the number of certificates is counted and not the number of students being awarded some form of Year 12 certification (Chapter 9 in COAG Reform Council, 2010).
5. For more information, see *http://nemp.otago.ac.nz/advertising/index.htm*.

References

ACT (Australian Capital Territory) (2010), *Annual Report 2009-2010*, ACT Department of Education and Training, Australian Capital Territory, Canberra.

Australian Bureau of Statistics (2006), *Deriving Measures of Engagement in Secondary Education from the National Schools Statistics Collection*, Research Paper, ABS, Canberra, *www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/DB6450FC8BFBDF78CA257243001B5EE2/$File/1351055016_dec%202006.pdf*.

Australian Government (2010a), *Department of Education, Employment and Workplace Relations – Annual Report 2009-10*, DEEWR, Australian Government, Canberra, *www.annualreport2010.deewr.gov.au/2010/Documents/DEEWRAnnualReport2009_10I.pdf*.

Australian Government (2010b), *Country Background Report for Australia*, prepared for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, Department of Education, Employment and Workplace Relations, Canberra, available from *www.oecd.org/edu/evaluationpolicy*.

Australian National Audit Office/CPA Australia (2008), *Monitoring and Reporting Financial and Non-financial Performance of Australian Government Organisations*, Issues Paper Series No. 5, February, *www.canberra.edu.au/corpgov-aps/pub/issuespaper5.pdf*.

Catholic Education Commission of Victoria Ltd (2010), *Annual Report 2009*, Catholic Education Commission of Victoria Ltd, East Melbourne, *http://web.cecv.catholic.edu.au/frameset.htm?page=positions*.

CITO (2008), *Onderwijs op peil? Een samenvattend overzicht van 20 jaar PPON*, Stichting Cito Instituut voor Toetsontwikkeling, Arnhem.

CITO (2009), *Peiling van de rekenvaardigheid en de taalvaardigheid in jaargroep 8 en jaargroep 4 in 2008 – Jaarlijks Peilingsonderzoek van het Onderwijsniveau*, Technische rapportage, Cito B.V., Arnhem.

COAG Reform Council (2010), *National Education Agreement: Performance Report for 2009*, COAG Reform Council, Sydney, *www.coagreformcouncil.gov.au/reports/docs/NEA_report_2009.pdf*.

Curriculum Corporation (2004), *National Year 6 Science Assessment Report 2003*, Curriculum Corporation; Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA), Victoria, Carlton South, *www.mceecdya.edu.au/verve/_resources/nat_year6_science_file.pdf*.

Curriculum Corporation (2008), *Melbourne Declaration on Educational Goals for Young Australians* – December 2008, Curriculum Corporation; Ministerial Council on Education, Employment, Training and Youth Affairs (MCEETYA), Victoria, Carlton South, *www.mceecdya.edu.au/verve/_resources/National_Declaration_on_the_Educational_Goals_for_Young_Australians.pdf.*

Department of Education (2010a), *Department of Education – Annual Report 2009-10*, Tasmania, Hobart, *www.education.tas.gov.au/annualreport/09-10.*

Department of Education (2010b), *Department of Education – Annual Report 2009-10*, The Government of Western Australia, East Perth, *www.det.wa.edu.au/education/AnnualReport/.*

Department of Education and Children's Services (2010), *Department of Education and Children's Services Annual Report 2009*, Government of South Australia, Adelaide.

Department of Education and Training (2010), *Department of Education and Training – Annual Report 2009-2010*, Department of Education and Training, Queensland Government, Brisbane.

Green, S. and T. Oates (2009), "Considering Alternatives to National Assessment Arrangements in England: Possibilities and Opportunities", *Educational Research*, Vol. 51, No. 2, June 2009, NFER, 229-245.

Lokan, J., L. Greenwood and J. Cresswell (2002), *15-up and Counting, Reading, Writing, Reasoning…. How Literate are Australia's Students?*, Australian Council for Education Research, Camberwell, Victoria, *www.acer.edu.au/documents/PISA_Report.pdf.*

Marks, G.N., J. McMillan, F.L. Jones and J. Ainley (2000), *The Measurement of Socioeconomic Status for the Reporting of Nationally Comparable Outcomes of Schooling*, Draft Report by the National Education Performance Monitoring Taskforce, Australian Council for Educational Research and Sociology Program, Research School of Social Sciences, Australian National University, *www.mceecdya.edu.au/verve/_resources/socioeconomicstatus_file.pdf.*

MCEECDYA (2010), *National Assessment Program – ICT Literacy Years 6 & 10 Report – 2008*, Ministerial Council for Education, Early Childhood Development and Youth Affairs, *www.mceecdya.edu.au/verve/_resources/NAP-ICTL_2008_report.pdf.*

MCEETYA (2000), *National Report on Schooling in Australia 2000*, Ministerial Council on Education, Employment, Training and Youth Affairs, *http://cms.curriculum.edu.au/anr2000/index.htm*.

MCEETYA (2007), *National Report on Schooling in Australia – Preliminary Paper: National Benchmark Results Reading Writing and Numeracy Years 3, 5 and 7*, Ministerial Council on Education, Employment, Training and Youth Affairs, *http://cms.curriculum.edu.au/anr2007/pdfs/2007benchmarks.pdf.*

MCEETYA (2008), *Measurement Framework for National Key Performance Measures*, Performance Measurement and Reporting Taskforce, Ministerial Council on Education, Employment, Training and Youth Affairs.

NCVER (National Centre for Vocational Education Research) (2009), *2008 VET in Schools Statistics*, A Report for the Ministerial Council for Education, Early Childhood Development and Youth Affairs (MCEECDYA) Secretariat, *www.mceecdya.edu.au/verve/_resources/NCVER_2008_VET_in_Schools_Report.pdf.*

New South Wales Department of Education and Training (2010), *Annual Report 2009* – NSW Department of Education and Training, State of New South Wales, Sydney, *www.det.nsw.edu.au/reports_stats/annual_reports*.

Ofsted (Office for Standards in Education) (2010), *The Annual Report of Her Majesty's Chief Inspector of Education, Children's Services and Skills 2009/10*, Ofsted, The Stationery Office, London.

Productivity Commission (2009), *A Quick Guide to the Productivity Commission*, Productivity Commission, Australian Government, *www.pc.gov.au/__data/assets/pdf_file/0005/64679/quick-guide-2009.pdf*.

Queensland Audit Office (2010), *Auditor-General of Queensland Report to Parliament No. 6 for 2010 – Using Student Information to Inform Teaching and Learning*, State of Queensland, Brisbane, *www.qao.qld.gov.au/downloadables/publications/auditor_general_reports/2010_Report_No.6.pdf*.

State of Victoria (2008), *Blueprint for Education and Early Childhood Development*, Department of Education and Early Childhood Development, Melbourne, *www.eduweb.vic.gov.au/edulibrary/public/commrel/policy/Blueprint2008/blueprint-final-2008.pdf*.

State of Victoria (2009), *Transparency and Accountability across all School Sectors in Victoria – Blueprint Implementation Paper*, Education Policy and Research Division, Office for Policy, Research and Innovation; Department of Education and Early Childhood Development, State of Victoria, Melbourne, *www.eduweb.vic.gov.au/edulibrary/public/commrel/policy/Blueprint2008/Blueprint_AllSectors.pdf*.

Steering Committee for the Review of Government Service Provision (2010), *Review of the Report on Government Services' Performance Indicator Framework – Report to the Steering Committee for the Review of Government Service Provision*, prepared by the Independent Reference Group, August 2010, *www.pc.gov.au/gsp/independent-reference-group-report*.

Thomson, S., L. De Bortoli, M. Nicholas, K. Hillman and S. Buckley (2011), *Challenges for Australian Education: Results from PISA 2009: The PISA 2009 Assessment of Students' Reading, Mathematical and Scientific Literacy*, Australian Council for Education Research, Camberwell, Victoria, *www.acer.edu.au/documents/PISA-2009-Report.pdf*.

United States Department of Education (2010), *Use of Education Data at the Local Level – From Accountability to Instructional Improvement*, U.S. Department of Education, Office of Planning, Evaluation and Policy Development, Washington D.C., *www2.ed.gov/rschstat/eval/tech/use-of-education-data/use-of-education-data.pdf*.

Wayman, J.C. (2005), "Involving Teachers in Data-Driven Decision Making: Using Computer Data Systems to Support Teacher Inquiry and Reflection", *Journal of Education for Students Placed at Risk*, 10, No. 3, 295–308.

Conclusions and recommendations

Education system context

Student learning outcomes are considerably above the OECD average but there is evidence of some decline

Student learning outcomes in Australia are very good by international standards even if there is evidence of some decline in the last decade. In 2009, achievement levels of Australian students in the OECD's Programme for International Student Assessment (PISA) were significantly above the OECD average in each of the assessment domains – reading, mathematics and science. However, trend analyses of PISA results have raised concerns about a decline in student learning outcomes – for example, Australia is among the five OECD countries for which student performance in reading declined significantly between 2000 and 2009. The variation in performance between high- and low-performing students in Australia was higher than the OECD average in reading and science, and similar to that found for the OECD as a whole in mathematics in PISA 2009. However, no statistically significant difference was observed in variation in student performance in reading between 2000 and 2009.

The national agenda for education reinforces the role of evaluation and assessment

In 2008 a major national agenda was established with a common framework for reform in education agreed between the Australian Government and the state and territory governments through the National Education Agreement (NEA). It developed from the National Productivity Agenda agreed by the Council of Australian Governments (COAG) and is supported by the Melbourne Declaration on Educational Goals for Young Australians, which articulates future directions and aspirations for Australian schooling. The main components of the national reform agenda are the development of the Australian Curriculum, the Aboriginal and Torres Strait Islander Education Action Plan, the National Partnerships, the National Assessment Program and the leadership of national-level entities such as the Australian Curriculum, Assessment and Reporting Authority (ACARA) and the Australian Institute of Teaching and School Leadership (AITSL). For the first time in Australia at the national level, the management of curriculum, assessment and reporting are brought together (through ACARA) and there is national leadership in the profession of teaching and school leadership (through AITSL). The NEA also brings an obligation to meet a common set of national school performance and reporting requirements. There is now a clearer framework of national expectations together with new national infrastructure and a firm commitment to improved transparency and accountability. In this context, the national agenda for education reinforces the role of evaluation and assessment as key tools to achieve quality and equity in education.

Strengths and challenges

Australia has a well-conceptualised evaluation and assessment framework but some articulations are not sufficiently developed

The national agenda for education has granted the opportunity to conceptualise evaluation and assessment at the national level through the development of goals, monitoring and reporting at the national level as well as mechanisms to articulate national objectives with jurisdiction-level goals and priorities. To the Review Team the overall evaluation and assessment framework appears as highly sophisticated and well conceptualised, especially at its top level (national and systemic levels). Particularly positive features include: the national educational goals as a solid reference point; strong capability at the national level to steer evaluation and assessment; a focus on student outcomes; a coherent system of assessments for learning; a structure to integrate accountability and improvement; and the commitment to transparency. However, there is a less clear articulation of ways for the national agenda to generate improvements in classroom practice through the assessment and evaluation procedures which are closer to the place of learning such as school evaluation, teacher appraisal and student formative assessment. This translates into a greater emphasis on the accountability function of evaluation and assessment as the improvement function is more articulated at the local level. The national education agenda has placed considerable investment in establishing national standards, national testing and reporting requirements while it provides considerably less direction and strategy on how to achieve the improvement function of evaluation and assessment. In addition, the Review Team noted a number of missing links, or underdeveloped articulations, between different elements of the overall evaluation and assessment framework. Examples include the alignment of teaching standards with teaching career structures; the articulation between teacher appraisal, school evaluation and school development; and the articulation between school self-evaluation and external school evaluation.

Striking the right balance between nationally-dictated policies and ability to meet local needs is a challenge

Given the current disparities of policy and practice in relation to evaluation and assessment procedures across Australia, a major challenge lies in determining what constitutes a desirable measure of consistency as against legitimate diversity. The nature of the national agenda for education is likely to be strengthened by greater consistency of evaluation and assessment procedures across jurisdictions but greater diversity offers more opportunities for innovation and adaptation to local needs. It is clear that much of what is required in student assessment, teacher appraisal, school evaluation and system evaluation is in place in aspects of current practice across jurisdictions and school sectors. The challenge is to articulate a national strategy for each of these evaluation and assessment components which builds on the best of current practice and continues to allow flexibility of approach within agreed parameters.

There is room to improve the integration of the non-governmental sector in the overall evaluation and assessment framework

The Melbourne Declaration places strong emphasis on the fact that Australian governments "commit to working with all school sectors" on all the key areas for schooling. While through the Schools Assistance Act 2008 non-government schools have an obligation to meet national school performance and reporting requirements similar to those which apply to government schools, the Review Team formed the impression that there is room to improve the integration of the non-governmental sector in the overall evaluation and assessment framework. The risk of a limited integration is that there is little guarantee that evaluation and assessment procedures in the Catholic and Independent sectors are sufficiently aligned with student learning objectives and educational targets at the national and systemic levels.

A coherent framework for the assessment of student learning is in place

A range of provisions for the assessment of student learning are established, including: the National Assessment Program Literacy and Numeracy (NAPLAN); triennial sampled-based assessments of ICT literacy, science literacy, and civics and citizenship; international assessments (*e.g.* PISA and TIMSS); A-E ratings; and senior secondary certificates and vocational education and training certificates. This set of assessments results in a coherent system of assessments of learning that potentially can provide a comprehensive picture of student performance relative to Australia's goals for student learning. That is, while NAPLAN and other periodic assessments provide a national barometer of performance necessarily on a limited set of standards (*i.e.* those that can be measured within limited testing time), A-E reporting requirements and secondary certificates provide a structure for linking accountability to a fuller set of national and/or educational jurisdictions' expectations for student learning. Performance on international measures enables policy makers and the public to monitor student progress over time against that in other countries.

NAPLAN results are credible and deemed useful but there are aspects to be improved

Most stakeholders find NAPLAN results a credible source of evidence. It is recognised that NAPLAN enables greater consistency, comparability and transferability of results across jurisdictions in a way that was not possible under the previous jurisdiction-based testing system. The use of a common scale is also valued as it provides significant information about the performance of, and growth in, individual student achievement. The trust placed in NAPLAN findings seems well justified from the perspective of the reliability and precision of reported scores as indicated by studies about the technical quality of the assessments. However, NAPLAN was developed and implemented prior to the introduction of the Australian Curriculum and thus may not be closely aligned with it. This limitation is being addressed with a review of the National Assessment Program subsequent to the release of the Australian Curriculum. A second alignment issue is that NAPLAN addresses a relative narrow range of learning goals relative to what parent, teacher, principal, and business representatives with whom we

spoke want for students. Clearly, this is not a problem unique to NAPLAN – there is a limit to what any time-limited, standardised test can address – but it is a potential concern if the system were to overemphasise NAPLAN results. While other components of the National Assessment Program may address other learning goals, the frequency and visibility of NAPLAN makes it a more important driver for Australia's educational system.

Summative student assessment is adequately supported by moderation processes and tools but there are some challenges in A-E reporting

The tools and resources developed by educational jurisdictions to support their schools and teachers' use of the A-E reporting scales, such as assessment tools and measurement standards linked to school curricula, appear very valuable to teachers. Similarly, procedures adopted by educational jurisdictions and particular schools for moderating A-E judgments and senior secondary assessments also are models for increasing the utility and consequences of assessment. However, a major challenge is to align A-E ratings to the Australian Curriculum, an undertaking which has now started under the leadership of ACARA. This will bring a national agreement on A-E definitions improving the current situation where A-E definitions differ across states and territories. Another challenge is to ensure that teachers develop capacity to assess against A-E ratings.

There is considerable reliance on teacher-based summative assessment but the emphasis on NAPLAN may "narrow" its use

There is a good focus on covering a broad range of evidence on student performance through teacher-based assessment in overall student summative assessments. Teachers' continuous classroom-based assessments are included in students' grades and typically contribute to the school-leaving certificate report. The practice of giving considerable weight to teacher-based assessment in student summative assessments is important. Nevertheless, there are indications that NAPLAN is becoming dominant in discussions around "student assessment". NAPLAN is given annually and school reporting makes it a highly visible assessment that is likely to send a strong signal to administrators, teachers and students about what is most important for teaching and learning. The risk is that the emphasis on NAPLAN may "narrow" teacher-based assessment.

Teacher registration processes are in place but there are some challenges to their implementation

Teacher registration processes are well established in Australian schools. They constitute a powerful quality assurance mechanism to ensure that every school in Australia is staffed with teachers with suitable qualifications who meet prescribed standards for teaching practice. At their initial level (provisional/graduate registration), they also provide a policy lever for setting entrance criteria for the teaching profession and, through the accreditation of initial teacher education programmes, strengthen the alignment between initial teacher education and the needs of schools. However, there are a number of aspects in implementation which deserve further policy attention. First, the

level of externality or external moderation in registration processes might not be adequate – processes are mostly school-based and the interpretation of standards is done at the local level with little moderation across schools. Second, registration standards do not fully reflect the complexity of teaching careers and the different levels of performance achievable with further experience. Third, as the maintenance of registration is essentially based on participation in professional development activities, there seems to be a weak link between registration's renewal and what teachers are actually doing in schools and what their students are learning.

Performance management processes provide a good basis for developmental teacher appraisal, which needs to become more systematic

Teacher appraisal as part of regular employer's performance management processes is expected to take place in Australian schools. In its current form, it has essentially an improvement function with the emphasis on evaluation for teacher development. This focus is suitable – it is intended to identify areas of improvement for individual teachers, and lead to the preparation of individual improvement plans (including professional development). However, there is evidence of great variation between schools in the way performance management is carried out, from a very light touch to it through to demanding and elaborate processes in some schools. Therefore there are no guarantees in Australian schools that performance management processes are addressing the real issues and complexities of teaching and learning, except in those schools where appraisal is well consolidated.

Teachers are trusted professionals with a high degree of autonomy but they have few opportunities for feedback

The Review Team formed the view that Australian teachers are generally perceived as trusted professionals among the different stakeholders. This is reflected in the extensive autonomy they benefit in the exercise of their duties. One of the results is that they are generally eager and willing to receive feedback. Teachers generally conveyed to the Review Team that they appreciated the time the school principal took to provide them with feedback and in general found classroom visits, where they occur, useful. However, Australian teachers have relatively few opportunities for professional feedback. The main opportunity to receive feedback on their practices is the annual performance review held with the school principal who tends to have limited time to engage properly in the coaching, monitoring and appraisal of teachers. Similarly, the interaction with experts of school review teams is infrequent and does not allow for a comprehensive review of teaching practices for individual teachers.

There is little alignment between teaching standards, registration processes and career structures

A problematic aspect of the teaching profession in Australia is that career structures are, in most jurisdictions, dissociated from teaching standards and registration processes. This translates into a detrimental separation between the definition of skills and competencies at different stages of the career (as reflected in teaching standards) and the roles and responsibilities of teachers in schools (as reflected in career structures). This is

problematic in a range of ways. In particular, it reduces the incentive for teachers to improve their competencies, and weakens the matching between teachers' levels of competence and the tasks which need to be performed in schools to improve student learning.

Accountability and transparency are well embedded as national principles guiding school evaluation but the role of school reviews is less well defined

The developing culture of school evaluation and improvement across Australia has already become particularly well established in a number of jurisdictions. The national policy environment has transparency and accountability as key planks in its improvement agenda. The language of accountability and transparency at the national, system and school levels is well aligned. School self-evaluation is an expectation and some form of external review mechanism is increasingly common. Test results, focusing on literacy and numeracy, are widely used to inform evaluation. However, there remains a need to clarify a number of vital issues relating to the relationship between the role of reviews in both accountability and improvement; the scope of reviews in relation to the emerging national agenda; the critical areas on which reviews should focus; the role and nature of externality; and the extent of transparency. Different jurisdictions have addressed mixtures of these issues in their own context but no clear national direction of travel has as yet emerged.

There are clear rules for school reporting and the principle of publishing performance data is established but there are potential undesired effects

The Principles and Protocols for Reporting on Schooling in Australia is a powerful document which makes clear the commitment to transparent accountability. The principles relate directly to data on student outcomes – publication of NAPLAN testing results on a school-by-school basis on the *My School* website – and information about the school context and resourcing – *e.g.* publication of school reports on the respective school website. The NAPLAN and associated *My School* website represent a powerful example of how a clear and well-articulated policy allied to determined and consistent leadership can bring about quick change. Hence, a key plank in the national policy agenda is the belief that measuring and publicising student outcomes on a comparative basis will lead schools to focus on taking the action necessary to improve their relative performance. Thus increased accountability and transparency will help drive improvement. There are, however, a number of possible undesired effects in placing too great reliance on this approach, not least the risk of a possible narrowing effect on the curriculum and wider achievement with an overemphasis on that which is assessed through the NAPLAN tests. There is also a danger that schools which perform satisfactorily may become complacent as the spotlight falls on those schools which perform least well comparatively.

External school reviews are well established but their focus needs to be better defined

Some form of external school review is widely in place across jurisdictions. The nature of externality is very much a matter for the jurisdiction concerned but the need for a view from outside the school itself seems to be common practice. A number of jurisdictions have recognised the need to engage reviewers who do not have any direct responsibilities associated with the school. However, there remains considerable variation in the focus of school reviews. Reviews need both to evaluate the outcomes being achieved and to identify the key factors which have influenced those outcomes such as the quality of teachers and the teaching process; the ethos of the school; leadership; and the capacity of the school to evaluate itself. It is important, therefore, to have a framework of criteria for evaluation which requires evidence about each of these factors and their relationship to the school's performance. Leadership in particular, a key factor in school effectiveness, does not seem to figure strongly in school evaluation frameworks.

Common reporting frameworks are well established

A core strength of the evaluation of education in Australia is the existence of clear standard frameworks both for reporting key performance measures and for general government sector reporting. The *Measurement Framework for National Key Performance Measures* establishes a set of measures to guide the development of nationally comparable data on aspects of performance critical to monitoring progress against national student learning objectives. This framework clearly presents the agreed measures and their source for each of the priority areas: literacy, numeracy, science literacy, civics and citizenship, ICT literacy, vocational education and training in schools, student participation, student attainment, student attendance, and in late 2010 was further refined to incorporate the full suite of agreed national key performance measures, including the COAG measures. The *Report on Government Services' Performance Indicator Framework* provides a common reporting basis for each government sector.

There are strong national monitoring tools but there remain challenges in some key measurement areas

A strong and stable set of national measures on education is established. National information to monitor education in Australia is largely compiled by the Australian Bureau of Statistics in collaboration with the different states and territories and the Australian Government. Along with ongoing work to improve the quality of national statistical indicators, the addition of comparable outcomes information from NAPLAN has significantly strengthened the set of national monitoring tools. While the significant progress that has been made in strengthening monitoring at the national level is commendable, there remain challenges in some key measurement areas. National measures are not available for all national goals and, in particular, data for some of the key student subgroups (*e.g.* Indigenous and socio-economically disadvantaged students) suffer in terms of coverage and quality at the national level. Also, the quality of some of the completion data is of concern.

There are strong procedures for system monitoring at the state and territory level but challenges remain in monitoring schools across sectors

Several jurisdictions have strengthened structures to monitor schooling over recent years, including through systemic approaches to school improvement, the formation of school networks and greater investment in performance analysis and reporting. Most jurisdictions offer software to aid schools in fulfilling their reporting responsibilities. All states and territories also have forms of testing in schools which are complementary to NAPLAN. Most jurisdictions also systematically collect qualitative feedback from the primary users of education, that is, the students and parents. There are also attempts to monitor student transitions after the completion of Year 12. A visible challenge is that the extent to which the systematic collection of information from the non-government sectors is in place varies considerably. The monitoring of non-government sectors is generally conducted via state or territory regulatory authorities, but reporting on their outcomes is still limited to a simple set of compliance statements and does not focus on performance.

There is an extensive use of results from the national monitoring system with some room for improvement

Many Australian government bodies make use of the results from the national monitoring system. The annual report from the Department of Education, Employment and Workplace Relations includes performance indicators that draw on results from NAPLAN and the national monitoring surveys, as well as enrolment and apparent retention rates. Results are also extensively reported in the *Report on Government Services* and the *National Report on Schooling in Australia*. Indicators in these reports are also extensively reported as part of state and territory government reporting. In general, results are systematically used to inform school improvement frameworks and policy development. A challenge is the increasing complexity of some outcome reporting which is often not understood by stakeholders such as parents or employers. There also seems to be potential to further exploit results from the national monitoring sample surveys in science, ICT, and civics and citizenship.

Policy recommendations

Establish national strategies for strengthening the linkages to classroom practice within the overall evaluation and assessment framework

Realising the full potential of the overall evaluation and assessment framework involves establishing strategies to strengthen the linkages to classroom practice, where the improvement of student learning takes place. A major step in this direction would be a national reflection about the nature and purpose of evaluation components such as school evaluation, teacher appraisal and student formative assessment within the overall education reform strategy and the best approaches for these evaluation components to improve classroom practices. This could lead to the establishment of a set of principles (or guidelines) on how to undertake or promote these activities in ways that support national student learning objectives. The principles should build on current best practice, align with the national policy agenda and respect traditions of Australian schooling.

Promote greater national consistency while giving room for local diversity

Greater consistency of evaluation and assessment practices across jurisdictions (and school sectors) would provide greater guarantees that such practices are aligned with national student learning objectives. An important first step might be to agree protocols between educational jurisdictions and the Australian Government for the design and implementation of given evaluation and assessment procedures. The protocols would involve the agreement of general principles for the operation of procedures such as school evaluation, teacher appraisal, student formative assessment or the evaluation of school leadership while allowing flexibility of approach within the agreed parameters to better meet local needs. The protocols should come along with clear goals, a range of tools and guidelines for implementation. They should permit better consistency of evaluation practices across educational jurisdictions while leaving sufficient room for local adaptation. This could imply requiring educational jurisdictions to develop action plans at the local level aligned with national protocols. The goals defined at the national and the jurisdiction level should be complementary in order to avoid conflicting messages to schools.

Improve the integration of the non-governmental sector in the overall evaluation and assessment framework

Evaluation and assessment practices in the Catholic and Independent sectors are very diverse and, with the exception of the reporting requirements which apply to all schools across Australia, display limited alignment with those in place in state and territory schools. As a result, in spite of well-consolidated practices in the non-government sector, there is limited guarantee that those practices are aligned with the national education agenda. Regarding evaluation and assessment procedures closer to the classroom (*e.g.* school evaluation, teacher appraisal), a possible solution to better integrate the non-governmental sector in the overall evaluation and assessment framework is for the non-government sector to be part of the protocol agreements suggested above to reach greater national consistency towards the national education agenda. This could become another requirement for non-government schools to receive public funding in a way similar to the reporting requirements.

Further develop some articulations within the overall evaluation and assessment framework

The process of developing an effective evaluation and assessment framework should give due attention to: achieving proper articulation between the different evaluation components (*e.g.* school evaluation and teacher appraisal); warranting the several elements within an evaluation component are sufficiently linked (*e.g.* teaching standards and teacher appraisal); and ensuring processes are in place to guarantee the consistent application of evaluation and assessment procedures (*e.g.* consistency of teachers' A-E ratings). For example, there are likely to be great benefits from the synergies between school evaluation and teacher appraisal. This indicates that school evaluation should comprise the monitoring of the quality of teaching and learning, possibly include the external validation of school-based processes for teacher appraisal, and school development processes should explore links to the evaluation of teaching practice.

Reinforce the assessment validity of NAPLAN and establish safeguards against an overemphasis on it

Further development of NAPLAN will need to address some important validity issues. First and foremost is NAPLAN's alignment with the Australian Curriculum and the extent to which NAPLAN is balanced in its representation of the depth and breadth of intended student learning goals (aspects currently being addressed by NAPLAN's developers). In addition, the specific purposes that NAPLAN may be expected to serve each bring implicit requirements for additional validity data, and requirements for serving some purposes may be at odds with others. For example, measures of student growth require tests that address a consistent set of targets over time and are vertically scaled, as NAPLAN is, which tends to narrow the breadth of content that can be assessed. Also, policy makers may want to consider safeguards and cross-checks to reduce the threat that the high visibility of NAPLAN results may encourage schools and educators to narrow the curriculum to the basic skills addressed by the current test, at the expense of knowledge and skills that are not annually assessed or reported at the individual student and school levels.

Strengthen teachers' capacity to assess student performance against the Australian Curriculum and to use student assessment data

In Australia's standards-based system, and in particular following the introduction of the Australian Curriculum, sound strategies to assess against the standards/curriculum are paramount. The current strategy for student assessment consists of a combination of NAPLAN and teacher-based assessments against the full range of curriculum goals (and reflected in A-E reporting). The latter implies a considerable investment on teacher capacity to assess against the standards, including specific training for teachers, the development of grading criteria and the strengthening of moderation processes within and across schools. This will be facilitated by the alignment of A-E ratings to the Australian Curriculum, an area of priority which is currently receiving attention through work led by ACARA. This work will bring the desirable consistency of A-E definitions across states and territories and will assure the proper link between teacher-based student assessment and the Australian Curriculum. Another priority is to develop teachers' capacity to use student assessment data, including that generated by NAPLAN, for the improvement of classroom instruction. This calls for the provision of formal training, possibly as a professional development option for teachers, on skills for analysing and interpreting student assessment data.

Maintain the centrality of teacher-based assessment while ensuring the diversity of assessment formats

The current prominence of NAPLAN within the student assessment framework requires particular care about not reducing the importance of teacher-based assessment. Several studies underline that teacher-based summative assessment has a greater potential to improve approaches to teaching and learning than external tests. However, it needs to be recognised that teacher-based assessments are often perceived as unreliable. This indicates that there is a case for combining teacher-based assessment with external assessment, which tends to be more reliable, especially when stakes for students are high.

Another approach is to develop on-demand assessments, where teachers can draw from a central bank of assessment tasks and ask students to take the assessment when they consider that they are ready.

Align teaching standards with a competency-based career structure for teachers

An important policy objective should be to align the definition of expected skills and competencies at different stages of the career (as reflected in teaching standards) and the tasks and responsibilities of teachers in schools (as reflected in career structures). This would strengthen the incentive for teachers to improve their competencies, and reinforce the matching between teachers' levels of competence and the tasks which need to be performed in schools to improve student learning. Such alignment can be achieved by developing teaching standards which reflect different levels of the teaching expertise needed in schools, and ensuring levels of teaching expertise match the key stages of the career structure.

Conceive teacher registration as career-progression evaluation

Given the alignment between teaching standards and the competency-based career structure for teachers, teacher registration can be conceived as career-progression evaluation. Career-progression evaluation would have as its main purposes holding teachers accountable for their practice, determining advancement in the career, and informing the professional development plan of the teacher. This approach would convey the message that reaching high standards of performance is the main road to career advancement in the profession. Appraisal for teacher registration, which is more summative in nature, needs to have a stronger component external to the school and more formal processes. It could be a mostly school-based process led by the school principal but it should include an element of externality such as an accredited external evaluator, typically a teacher from another school with expertise in the same area as the teacher being appraised.

Perform developmental evaluation through teacher appraisal as part of performance management processes

Teacher appraisal as part of performance management processes should be conceived as developmental evaluation, *i.e.* the main process through which the improvement function of teacher appraisal is achieved. It would retain its current character but school-based processes for developmental evaluation would need to be strengthened and validated externally. Given that there are risks of bringing together both the accountability and improvement functions in a single teacher appraisal process, it is recommended that teacher appraisal as part of performance management processes is conceived as predominantly for improvement while teacher appraisal for registration performs a primarily accountability function. The developmental evaluation would be an internal process carried out by line managers, senior peers, and the school principal. The reference standards would be the teaching standards but with school-based indicators and criteria. This appraisal should also take account of the school objectives and activity plan.

The main outcome would be feedback on teaching performance as well as on the overall contribution to the school which would lead to a plan for professional development.

Clearly establish the fundamental purpose of external school evaluation

As part of a general agenda, the fundamental purpose of external school evaluation needs to be more clearly and consistently understood. School evaluation can be part of the strategy to bring about general improvement across all schools or, more narrowly, it can focus on "failing schools". The approach adopted depends on the underlying policy agenda and the evidence about the performance of the school system as a whole. However, a rigorous but constructive approach to evaluation is seen by many countries as a means of driving improvement while also satisfying the needs of accountability.

Strengthen the alignment between self-evaluation and external evaluation, and ensure a broad scope for external school evaluation

Moves towards achieving a much closer alignment between self-evaluation and external evaluation could prove beneficial. The central requirement is that internal evaluation and external evaluation use common criteria and share a common language of quality. Where this is not the case, the school can be pulled in a variety of different directions with no strong evidence base to determine priorities. The criteria can be expressed in different ways but they should focus on those areas which are known to be critical factors in school quality. Another policy priority relates to the nature of external evaluation itself. There is no single, prevailing approach to who should be engaged in external evaluation but there is a need to establish clear expectations about externality which will apply across jurisdictions. The scope and frequency of external review are also important issues. The implementation of the broadening Australian Curriculum suggests a more general focus than that which a "failing schools" agenda might imply. For these reasons, developing policy on school evaluation in Australia should seek to use its potential to challenge complacency and provide evidence about progress on a broad front.

Publish externally validated school evaluation reports to complement the publication of national test data

Given the publication of comparative national test data, there remains a strong case to provide complementary evaluative information which broadens the base of evidence and provides more explanation of the factors which have influenced performance. Arguably, testing can only provide a *post hoc* evaluation of performance but good school evaluation is more proactive and should help to identify those factors which are influencing performance at an earlier stage. Consideration should therefore be given to not only continuing to refine and extend the content of the *My School* website but to include direct links to school reports which are validated by external involvement, are more comprehensive in their scope, look inside the "black box" of the working of the school and set a clear improvement agenda.

Continue and prioritise efforts to meet information needs for national monitoring

The Review Team endorses the two priority areas identified by the COAG Reform Council to improve performance reporting: "achievement of Indigenous students and students from low socio-economic backgrounds" and "reporting of change over time". The immediate priority for meeting information needs to adequately monitor progress towards national goals is to strengthen the information systems regarding student socio-economic and Indigenous status. The quality of socio-economic background data, in particular, proves inadequate for monitoring progress on several key indicators.

Further exploit results from jurisdiction and national monitoring systems for systemic school improvement

States and territories should continue efforts to strengthen monitoring structures, in part by further exploiting the analysis of results from local information systems and the national monitoring system, and importantly by ensuring adequate monitoring and follow-up on priority areas (*e.g.* underperforming schools) and the impact of departmental interventions.

Support and promote greater monitoring in the non-government sector

There may be ways to more efficiently meet state and territory government responsibilities for "timely, consistent and comparable reporting" in all school sectors. Strengthened administrative data collections would make a key contribution to this end. Another possibility is for states and territories to establish common performance summary reports for schools in all sectors. Another possibility would be to include the monitoring information on non-government sectors as part of the annual government education department reporting.

Annex A: The OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes

The OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes is designed to respond to the strong interest in evaluation and assessment issues evident at national and international levels. It provides a description of design, implementation and use of assessment and evaluation procedures in countries; analyses strengths and weaknesses of different approaches; and provides recommendations for improvement. The Review looks at the various components of assessment and evaluation frameworks that countries use with the objective of improving student outcomes. These include student assessment, teacher appraisal, school evaluation and system evaluation. The Review focuses on primary and secondary education.[1]

The overall purpose is to explore how systems of evaluation and assessment can be used to improve the quality, equity and efficiency of school education.[2] The overarching policy question is "How can assessment and evaluation policies work together more effectively to improve student outcomes in primary and secondary schools?" The Review further concentrates on five key issues for analysis: *(i)* Designing a systemic framework for evaluation and assessment; *(ii)* Ensuring the effectiveness of evaluation and assessment procedures; *(iii)* Developing competencies for evaluation and for using feedback; *(iv)* Making the best use of evaluation results; and *(v)* Implementing evaluation and assessment policies.

Twenty-four countries are actively engaged in the Review. These cover a wide range of economic and social contexts, and among them they illustrate quite different approaches to evaluation and assessment in school systems. This will allow a comparative perspective on key policy issues. These countries prepare a detailed background report, following a standard set of guidelines. Countries can also opt for a detailed Review, undertaken by a team consisting of members of the OECD Secretariat and external experts. Eleven OECD countries have opted for a Country Review. The final comparative report from the OECD Review, bringing together lessons from all countries, will be completed in 2012.

The project is overseen by the Group of National Experts on Evaluation and Assessment, which was established as a subsidiary body of the OECD Education Policy Committee in order to guide the methods, timing and principles of the Review. More details are available from the website dedicated to the Review: *www.oecd.org/edu/evaluationpolicy*.

Notes

1. The scope of the Review does not include early childhood education and care, apprenticeships within vocational education and training, and adult education.
2. The project's purposes and scope are detailed in OECD 2009 document entitled "OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Design and Implementation Plan for the Review", which is available from the project website *www.oecd.org/edu/evaluationpolicy*.

Annex B: Visit itinerary

Monday 21 June, Melbourne

08:30 – 09:30 Evaluation and Assessment Frameworks in the Victorian school education portfolio

- Department of Education and Early Childhood Development of Victoria
  - Office for Policy, Research and Innovation
  - Office for Children and Portfolio Coordination
- Independent Schools Victoria
- Catholic Education Commission Victoria
- Victoria Registration and Qualification Authority
- Victoria Institute of Teaching
- Victoria Curriculum and Assessment Authority

09:30 – 10:30 Government School Focus (School-level assessment and evaluation, teacher evaluation)

- Department of Education and Early Childhood Development of Victoria
  - Office for Policy, Research and Innovation
  - Office for Children and Portfolio Coordination

  School Accountability and Improvement Framework; Victorian School Performance Summaries, including intake-adjusted measures; Teacher Performance and Development Process; Performance and Development Culture; e5 Instructional Model; Principal Performance and Development Process; Bastow Institute of Educational Leadership; A to E reporting; English Online Interview.

10:30 – 11:30 Government School Stakeholders

- Parents Victoria
- Victorian Council of state School Organisations
- Association of School Councils in Victoria
- Australian Education Union Victoria

11:30 – 12:30 Catholic Education Commission

12:30 – 13:30 Independent Schools Victoria (ISV) and principals of independent schools

14:00 – 16:00 School visit: Ringwood Secondary College

- School leadership team
- Meeting with a group of teachers
- Meeting with a group of students

16:30 – 18:00 Regional Office visit: Eastern Metro Region

Tuesday 22 June, Melbourne

08:30 – 10:15 Independent School visit: Gilson College

- School leadership team
- Meeting with a group of teachers
- Meeting with a group of students

11:00 – 11:45 Victorian Auditor-General's office, Director Performance Audit

11:45 – 12:30 Productivity Commission

- Assistant Commissioner and Head of Secretariat for the Steering Committee for the Review of Government Service Provision
- Research Management, Secretariat for the Steering Committee for the Review of Government Service Provision

12:30 – 13:00 Australian Curriculum Assessment and Reporting Authority (ACARA), Chair of Board

13:00 – 14:00 Working lunch with ACARA and AITSL

14:00 – 14:30 Australian Institute for Teaching and School Leadership (AITSL), Chair of Board

14:30 – 15:15 Australian Council of Deans of Education (ACDE)

15:30 – 17:00 Research Seminar

- Australian Council for Educational Research (ACER) (Geoff Masters)
- Grattan Institute (Ben Jensen, Julian Reichl, Katherine Molyneux)
- Melbourne University (Patrick Griffin, Esther Care, Suzanne Rice)
- National Centre for Vocational Education Research (Nhi Nguyen)

Wednesday 23 June, Canberra

09:00 – 09:30 Department of Education, Employment and Workplace Relations

- Deputy Secretary Schooling Cluster

10:00 – 10:30 Deputy Prime Minister's Office, Parliament House

- Deputy Chief of Staff

11:00 – 11:30 Department of Education, Employment and Workplace Relations (DEEWR)

Schooling Cluster Executive Roundtable
High-level discussion of national initiatives and DEEWR's role
(National Education Agreement, National Partnerships, Digital Education Revolution, Funding Review, etc.)

- Group Manager, National Curriculum, Assessment and Reporting
- A/g Group Manager, Lifting Educational Outcomes
- Group Manager, Infrastructure and Funding
- Group Manager, Digital Youth and Transitions

11:30 – 12:15 Department of Education, Employment and Workplace Relations

Lifting Educational Outcomes Group: access, equality and Indigenous education issues as well as national developments in teaching
A/g Group Manager, Lifting Educational Outcomes

- Branch Manager, Inclusive Education
- Branch Manager, Indigenous Education
- Branch Manager, School and Student Support
- Branch Manager, Teaching Reforms Branch

12:15 – 13:00 Department of Education, Employment and Workplace Relations

National Curriculum, Assessment and Reporting Group
Group Manager, National Curriculum, Assessment and Reporting

- Branch Manager, Reporting and Accountability
- Branch Manager, National Curriculum

13:00 – 14:00 Industry groups

Australian Chamber of Commerce and Industry

- Director of Employment, Education and Training

Australian Industry Group (based in New South Wales)

- National Manager, Policy and Projects, Education and Training

14:00 – 15:00 National Non-government Organisations

Independent Schools Council Australia (ISCA)

- Deputy Executive Director
- Manager, Policy Analysis & Research

National Catholic Education Commission (NCEC)

- Executive Officer

15:00 – 15:45 Australian Education Union

- Deputy Federal Secretary
- Federal Research Officer

15:45 – 16:30 Independent Education Union

- Federal Secretary

16:30 – 17:15 Australian Council of state Schools Organisation

- National Projects Manager

Thursday 24 June, Brisbane (half of the team)

09:00 – 11:30 Government School visit: Durack state School

- School leadership team
- Meeting with a group of teachers
- Meeting with a group of students

12:30 – 15:00 Government School visit: Benowa state High School

- School leadership team
- Meeting with a group of teachers
- Meeting with a group of students

Thursday 24 June, Perth (half of the team)

09:00 – 11:30 Public School visit: Willetton Senior High School

- School leadership team and representatives of Burrendah PS and Castlereagh School
- Meeting with a group of teachers
- Meeting with a group of students

12:00 – 14:00 Department of Education of Western Australia

- Independent Public Schools and the Expert Review Group, Directors Schools, Evaluation and Accountability, Curriculum, HR personnel with responsibilities and expertise in assessment, performance and evaluation (students, teachers, schools, systems)

14:00 – 14:45 Aboriginal Stakeholders

- Aboriginal Education and Training Council
- Catholic Education Aboriginal Committee (Catholic Education Commission)
- Department of Education representatives
- Association of Independent Schools representatives
- Aboriginal Independent Community Schools Support Unit

14:45 – 15:30 Parent Groups

- Western Australia Council of state Schools Organisation
- Parents and Friends' Federation of Western Australia
- Parent Advisory Committee (Catholic Education Commission)

15:30 – 16:15 Unions, State School Teachers' Union of Western Australia

16:15 – 17:30 Principal Groups

- Western Australia Primary Principals' Association
- Western Australia Secondary School Executives' Association
- Western Australia District High School Administrators' Association
- Western Australia Education Support Principals and Administrators' Association
- Catholic Secondary Principals' Association of Western Australia
- Catholic Primary Principals' Association of Western Australia
- Independent Primary School Heads of Australia (Western Australia)
- Association of Heads of Independent Schools (Western Australia)
- Australian Special Education Principals' Association

Friday 25 June, Brisbane (half of the team)

08:30 – 10:30 Evaluation and Assessment Frameworks in Queensland

- Department of Education and Training of Queensland
  - A/g Assistant Director-General, Teaching and Learning
  - Executive Director School Improvement, Metropolitan
  - Director General, Human Resources
  - Assistant Director-General, Corporate Strategy and Performance
  - Director, School Performance Policy
  - Executive Director, Education Strategic Policy
  - Executive Director, School Improvement South East
- Independent Schools Queensland
- Queensland Catholic Education Commission

10:30 – 12:30 Principal Associations

- Queensland Association of state School Principals
- Joint Council of Queensland Teacher Associations
- Queensland Secondary Principals Association
- Queensland state P-10/12 School Administrators Associations
- Association of Heads of Independent Schools of Australia (Queensland)
- Association of Principals of Catholic Secondary Schools of Queensland
- Association for Special Education Administrators
- Queensland Catholic Primary Principal Association

12:30 – 14:00 Statutory Authorities

- Queensland College of Teachers
- Queensland Studies Authority
- Non state Accreditation Board

14:00 – 15:00 Unions

- Queensland Teachers Union
- Queensland Independent Education Union

15:00 – 16:00 Indigenous Consultative Groups

- Queensland Indigenous Education Consultative Committee
- Indigenous Education Leadership Institute, QUT

16:00 – 17:15 Parent Groups

- The Federation of Parents & Friends Association of Catholic Schools in Queensland
- Isolated Children's Parents' Association of Australia
- Queensland Independent Schools Parents Council
- Queensland Council of Parents and Citizens Association

Friday 25 June, Perth (half of the team)

08:30 – 11:00 Independent School visit: Scotch College

- School leadership team
- Meeting with a group of teachers
- Meeting with a group of students

11:30 – 12:30 Statutory Agencies

- Curriculum Council of Western Australia
- Department of Education Services
- Western Australia College of Teaching

12:30 – 14:00 Non Government School Sector

- Catholic Education Office
- Association of Independent Schools of Western Australia

14:00 – 14:45 Sector heads

- Department of Education
- Catholic Education Office
- Association of Independent Schools of Western Australia

Sunday 27 June, Sydney

Review Team meetings

Monday 28 June, Sydney

8:30 – 10:30 New South Wales (NSW) Department of Education and Training

Executive Team / Student Engagement and Program Evaluation Bureau / Educational Measurement and School Accountability Directorate / Strategic Planning

- Deputy Director-General, Strategic Planning and Regulation
- Acting General Manager, Planning and Innovation
- R/General Manager, External Relations Policy
- Deputy Director-General, Schools
- General Manager, Learning and Development
- General Manager, Access and Equity
- Director, Professional Learning and Leadership Development
- Relieving Director, Educational Measurement and School Accountability
- Senior Manager, Strategic Coordination
- Senior Manager, Student Engagement and Program Evaluation Bureau

10:30 –11:30 Non-government School Organisations

- NSW Catholic Education Commission (NSW CEC)
- NSW Association of Independent Schools (AISNSW)

11:30 – 12:15 NSW Board of Studies

- Chief Executive
- Director, Examinations and Credentials

12:15 – 13:00 NSW Institute of Teachers

13:30 – 16:00 School visit: Dulwich High School of Visual Arts and Design

- School leadership team and School Education Director, Sydney Inner City
- Meeting with a group of teachers
- Meeting with a group of students

16:30 – 18:00 Visit to Regional Office: Sydney Region Office

- Regional Director, Sydney
- School Education Director, Sydney Inner City
- Professional Support Officer, Sydney Region

Tuesday 29 June, Sydney

09:00 – 11:15 Non-government (Catholic) School visit: Trinity College Auburn

- School leadership team and Catholic Education Office
- Meeting with a group of teachers
- Meeting with a group of students

12:00 – 12:45 Indigenous Groups

- NSW Aboriginal Education Consultative Group Inc (AECG)
- Aboriginal Education and Training Directorate (DET)

12:45 – 13:30 Special Needs Groups

- Australian Association of Special Education (AASE)
- National Independent Special Schools Association (NISSA)
- Children with Disability Australia (CDA)
- Australian Advisory Board on Autism Spectrum Disorders

13:30 – 14:00 National Professional Teaching Association

- Australian Professional Teaching Association (APTA)
- Australian College of Educators (based in ACT)

14:00 – 14:45 State-level Parents Groups

- NSW Parents' Council Inc
- Federation of Parents and Citizens' Associations of NSW
- Council of Catholic School Parents NSW/ACT

14:45 – 15:30 State-level Teachers Groups

- NSW Teachers Federation
- NSW/ACT Independent Education Union

15:30 – 16:30 National Principals Groups

- Australian Primary Principals Association (APPA)
- Australian Secondary Principals Association (ASPA)
- Australian Heads of Independent Schools Australia
- Catholic Secondary Principals Australia (CaSPA)

16:30 – 17:15 National Professional Association

- Australian Council for Educational Leaders (ACEL)

17:15 – 18:00 State-level Principals Groups

- NSW Secondary Principals' Council (NSWSPC)
- NSW Primary Principals' Association (NSWPPA)
- Association of Catholic School Principals NSW
- Independent Primary School Heads of Australia (NSW Branch)

Wednesday 30 June, Sydney

09:00 – 09:45 Australian Curriculum, Assessment and Reporting Authority (ACARA)

- General Manager, Curriculum and Deputy CEO
- A/General Manager, Reporting
- A/General Manager, Assessment

09:45– 10:30 Australian Capital Territory (ACT) Department of Education and Training

- Director, Measurement Monitoring and Reporting
- Senior Manager, Measurement

10:30 – 11:15 South Australia (SA) Department of Education and Children's Services

- Manager, Improvement and Accountability, Quality, Improvement and Effectiveness Unit

11:15 – 12:00 NSW Department of Education and Training

SMART (School Measurement, Assessment and Reporting Toolkit) Presentation

- Relieving Director, Educational Measurement and School Accountability

12:00 – 13:30 Oral Report by Review Team with preliminary conclusions

Annex C: Composition of the Review Team

Graham Donaldson was Her Majesty's Senior Chief Inspector of Education from 2002 to 2010 in Scotland. In that role, he was Chief Executive of HM Inspectorate of Education and Chief Professional Adviser to the Scottish Government on all aspects of education outside the university sector. Graham began his teaching career in 1970 and taught in schools in Glasgow and Dunbartonshire. He worked as a Curriculum Evaluator for the Consultative Committee on the Curriculum. During this period, he was seconded to BP to review links between education and industry. His report, *Industry and Scottish Schools*, was published in 1981. He became an HM Inspector in 1983. Graham is the current President of the Standing International Conference of Inspectorates (SICI) which has 29 member inspectorates from across Europe. Following his retirement from HMIE, Graham has been asked by Scottish Government to undertake a national review of teacher education in Scotland. Graham was awarded a CB for his services to education in the 2009 Queen's Birthday Honours List.

Joan Herman, a United States national, is Director of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST) at the University of California – Los Angeles. Her research has explored the effects of testing on schools, the design of assessment systems to support school planning and instructional improvement, and the quality of teachers' formative assessment practices in mathematics and science. She also has wide experience as an evaluator of school reform. A former teacher and school board member, she is past president of the California Educational Research Association and has held a variety of elected leadership positions in the American Educational Research Association (AERA). Among her current involvements, she is editor of the research journal *Educational Assessment*, member of the Joint Committee for the Revision of the Standards for Educational and Psychological Measurement, member at large for AERA, and chairs the Board of Education for *Para Los Niños*.

Paulo Santiago, a Portuguese national, is a Senior Analyst in the OECD Directorate for Education, where he has been since 2000. He is currently the co-ordinator of the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes. He has previously assumed responsibility for two major cross-country reviews, each with the participation of over twenty countries: a review of teacher policy (between 2002 and 2005, leading to the OECD publication *Teachers Matter*) and the thematic review of tertiary education (between 2005 and 2008, leading to the OECD publication *Tertiary Education for the Knowledge Society*). He has also led reviews of teacher policy and tertiary education policy in several countries. He holds a PhD in Economics from Northwestern University, United States, where he also lectured. With a background in the economics of education, he specialises in education policy analysis. He co-ordinated this Review and acted as the *Rapporteur* for the Review Team.

Claire Shewbridge, a British national, is an Analyst in the OECD Directorate for Education and is currently working for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes. She most recently worked on the OECD Review on Migrant Education working on country-specific analysis for the Netherlands, Austria and Norway and co-authored the OECD report *Closing the Gap for Immigrant Students* (2010). For five years, Claire co-ordinated the PISA thematic report series. She also led analysis of student attitudes towards science learning and the environment in the PISA 2006 survey. Her earlier statistical work with the OECD included educational enrolment, graduation and financial statistics published in *Education at a Glance*, labour force survey statistics published in the *OECD Employment Outlook* and financial statistics in the OECD's Development Assistance Committee.

Annex D: Comparative indicators on evaluation and assessment

	Australia	Country Average[1]	Australia's Rank[2]
EDUCATIONAL ATTAINMENT Source: Education at a Glance (OECD, 2010a)[3]			
% of population that has attained at least upper secondary education, by age group (excluding ISCED 3C short programmes)[4] (2008)			
Ages 25-64	70	71	=17/30
Ages 25-34	82	80	=19/30
Ages 35-44	73	75	21/30
Ages 45-54	66	68	18/30
Ages 55-64	55	58	=19/30
% of population that has attained tertiary education, by age group (2008)			
Ages 25-64	36	28	=7/31
Ages 25-34	42	35	=8/31
Ages 35-44	38	29	=7/31
Ages 45-54	33	25	6/31
Ages 55-64	28	20	=5/31
Upper secondary graduation rates (2008) % of upper secondary graduates (first-time graduation) to the population at the typical age of graduation[5]	m	80	m
STUDENT PERFORMANCE			
Mean performance in PISA (Programme for International Student Assessment) (15-year-olds) Source: PISA 2009 Results (OECD, 2010c)[3]			
Reading literacy	515	493	6/34
Mathematics literacy	514	496	9/34
Science literacy	527	501	7/34
SCHOOL SYSTEM EXPENDITURE Source: Education at a Glance (OECD, 2010a)[3]			
Expenditure on primary, secondary and post-secondary non-tertiary institutions as a % of GDP, from public and private sources			
1995	3.5	~	18/26
2000	3.6	~	=13/29
2007	3.5	3.6	=17/29
Public expenditure on primary, secondary and post-secondary non-tertiary education as a % of total public expenditure (2008)[6]	10.3	9.0	8/29
Total expenditure on primary, secondary and post-secondary non-tertiary education from public sources (2007) (%)	81.1	90.3	22/25
Annual expenditure per student by educational institutions, (2007) (USD)[7]			
Primary	6498	6741	16/28
Lower secondary	8967	7598	13/26
Upper secondary	8639	8746	16/26
All secondary	8840	8267	13/28
Change in expenditure per student by educational institutions, primary, secondary and post-secondary non-tertiary education, index of change between 1995, 2000 and 2007 (2000 = 100)			
1995	85	88	14/22
2007	116	125	16/27
Current expenditure – composition, primary, secondary and post-secondary non-tertiary education (2007)[8]			
Compensation of teachers	62.0	63.8	11/20
Compensation of other staff	15.3	14.9	9/20
Compensation of all staff	77.2	79.2	18/28
Other current expenditure	22.8	20.8	11/28

	Australia	Country Average[1]	Australia's Rank[2]
SCHOOL STAFF NUMBERS			
Ratio of students to teaching staff (2008)			
Source: Education at a Glance (OECD, 2010a)[3,9]			
Primary	15.8	16.4	=13/27
Lower Secondary	m	13.7	m
Upper Secondary	m	13.5	m
All Secondary[10]	12.0	13.7	=17/29
CHARACTERISTICS OF THE TEACHER WORKFORCE (lower secondary education, 2007-08)			
Source: TALIS (OECD, 2009b)[11]			
Age distribution of teachers			
Teachers aged under 25 years	4.5	3.0	5/23
Teachers aged 25-29 years	13.7	12.1	7/23
Teachers aged 30-39 years	22.6	28.0	19/23
Teachers aged 40-49 years	26.5	29.6	16/23
Teachers aged 50-59 years	28.9	23.5	6/23
Teachers aged 60 years and more	3.8	3.9	9/23
Gender distribution of teachers (% of females)	59.2	69.3	19/23
Teachers' educational attainment[4]			
% of teachers who completed an ISCED 5A qualification or higher	98.7	83.7	4/23
Employment status of teachers			
% of teachers permanently employed	86.8	84.5	=9/23
TEACHER SALARIES in public institution			
Source: Education at a Glance (OECD, 2010a)[3]			
Annual teacher salaries (2008)[7]			
Primary – starting salary (USD)	33153	28949	8/29
Primary – 15 years experience (USD)	46096	39426	8/29
Primary – top of scale (USD)	46096	48022	17/29
Primary – ratio of salary after 15 years experience to GDP per capita	1.25	1.16	12/29
Lower secondary – starting salary (USD)	33336	30750	8/29
Lower secondary – 15 years experience (USD)	46908	41927	9/29
Lower secondary – top of scale (USD)	46908	50649	17/29
Lower secondary – ratio of salary after 15 years experience to GDP per capita	1.27	1.22	10/29
Upper secondary – starting salary (USD)	33336	32563	10/28
Upper secondary – 15 years experience (USD)	46908	45850	14/28
Upper secondary – top of scale (USD)	46908	54717	18/28
Upper secondary – ratio of salary after 15 years experience to GDP per capita	1.27	1.29	13/28
Number of years from starting to top salary (lower secondary education) (2008)	9	24	24/27
Decisions on payments for teachers in public schools (2008)[12]			
Criteria for base salary and additional payments awarded to teachers in public institutions			
● Base salary/■ Additional yearly payment /Δ Additional incidental payment			
Years of experience as a teacher	●	●29 ■9 Δ8	
Management responsibilities in addition to teaching duties	●	●12 ■18 Δ7	
Teaching more classes or hours than required by full-time contract	-	●2 ■10 Δ17	
Special tasks (career guidance or counselling)	-	●4 ■13 Δ11	
Teaching in a disadvantaged, remote or high cost area (location allowance)	●	●9 ■18 Δ4	
Special activities (*e.g.* sports and drama clubs, homework clubs, summer schools etc.)	-	●1 ■8 Δ12	
Teaching students with special educational needs (in regular schools)	■	●9 ■11 Δ5	
Teaching courses in a particular field	-	●5 ■8 Δ4	
Holding an initial educational qualification higher than the minimum qualification required to enter the teaching profession	●	●18 ■9 Δ5	
Holding a higher than minimum level of teacher certification or training obtained during professional life	●	●15 ■11 Δ3	
Outstanding performance in teaching	-	●5 ■9 Δ8	
Successful completion of professional development activities	-	●10 ■7 Δ4	
Reaching high scores in the qualification examination	-	●4 ■3 Δ3	
Holding an educational qualification in multiple subjects	-	●3 ■4 Δ3	
Family status (married, number of children)	■	●2 ■8 Δ1	
Age (independent of years of teaching experience)	-	●4 ■3 Δ1	
Other	-	●1 ■8 Δ2	

	Australia	Country Average[1]	Australia's Rank[2]
TEACHER PROFESSIONAL DEVELOPMENT (lower secondary education) Source: TALIS (OECD, 2009b)[11]			
Teacher participation in professional development (2007-08)			
% of teachers who undertook some prof. development in the previous 18 months	96.7	88.5	3/23
Average days of professional development across all teachers	8.7	15.3	18/23
Average days of professional development among those who received some	9.0	17.3	19/23
Average % of professional development days taken that were compulsory	47.3	51.0	11/23
Types of professional development undertaken by teachers (2007-08)			
Courses and workshops	90.6	81.2	6/23
Education conferences and seminars	64.0	48.9	5/23
Qualification programmes	11.7	24.5	20/23
Observation visits to other schools	22.2	27.6	12/23
Professional development network	60.1	40.0	4/23
Individual and collaborative research	36.6	35.4	12/23
Mentoring and peer observation	48.6	34.9	4/23
Reading professional literature	82.4	77.7	11/23
Informal dialogue to improve teaching	93.7	92.6	12/23
Impact of different types of professional development undertaken by teachers (2007-08) % of teachers reporting that the professional development undertaken had a moderate or high impact upon their development as a teacher			
Courses and workshops	78.5	80.6	16/23
Education conferences and seminars	67.6	73.9	21/23
Qualification programmes	78.6	87.2	21/23
Observation visits to other schools	72.2	74.9	14/23
Professional development network	73.5	80.2	19/23
Individual and collaborative research	85.8	89.3	20/23
Mentoring and peer observation	72.5	77.6	18/23
Reading professional literature	66.4	82.8	22/23
Informal dialogue to improve teaching	86.0	86.7	13/23
Teachers' high professional development needs (2007-08) % of teachers indicating they have a 'high level of need' for professional development in the following areas			
Content and performance standards	8.3	16.0	17/23
Student assessment practices	7.5	15.7	19/23
Classroom management	5.2	13.3	21/23
Subject field	5.0	17.0	=19/23
Instructional practices	3.6	17.1	23/23
ICT teaching skills	17.8	24.7	18/23
Teaching special learning needs students	15.1	31.3	22/23
Student discipline and behaviour problems	6.6	21.4	23/23
School management and administration	5.9	9.7	16/23
Teaching in a multicultural setting	4.0	13.9	22/23
Student counselling	7.3	16.7	22/23
TEACHER PERCEPTION OF SELF-EFFICACY (lower secondary education) Source: TALIS (OECD, 2009b)[11]			
% of teachers who 'strongly agree' or 'agree' with the statement "Teachers feel that they are making a significant educational difference" (2007-08)	93.7	92.3	12/23
% of teachers who 'strongly agree' or 'agree' with the statement "Teachers feel that when they try really hard, they can make progress with even the most difficult and unmotivated students" (2007-08)	87.0	82.7	7/23
SYSTEM EVALUATION			
Potential subjects of assessment at national examinations[13] (lower secondary education) (2006) Source: Education at a Glance (OECD, 2008)[3,14]			
National examinations exist (Yes/No)	No	8/25	
Mathematics	a	9/9	
Science	a	7/9	
National language or language of instruction	a	9/9	
Other subjects	a	8/9	
Compulsory for schools to administer national examinations (Yes/No)	a	7/9	
Year/Grade of national examination	a	9.2	

	Australia	Country Average[1]	Australia's Rank[2]
Potential subjects of assessment at national periodical assessment[15] (lower secondary education) (2006) Source: Education at a Glance (OECD, 2008)[3,14]			
National periodical assessments (Yes/No)[16]	Yes	14/25	
Mathematics	Yes	12/13	
Science	No	5/13	
National language or language of instruction	Yes	12/13	
Other subjects	No	6/12	
Compulsory for school to administer national assessment (Yes/No)	Yes	10/13	
Year/Grade of national assessment	7	-	
Possible influence of national examinations (lower secondary education) (2006) Source: Education at a Glance (OECD, 2008)[3] None/Low/Moderate/High[17]			
Performance feedback to the school	a	None:2 Low:1 Moderate:1 High:3	
Performance appraisal of the school management	a	None:4 Low:1 Moderate:1 High:1	
Performance appraisal of individual teachers	a	None:4 Low:2 Moderate:0 H igh:1	
The school budget	a	None:7 Low:1 Moderate:0 High:0	
The provision of another financial reward or sanction	a	None:7 Low:1 Moderate:0 High:0	
The assistance provided to teachers to improve their teaching skills	a	None:3 Low:0 Moderate:3 High:0	
Remuneration and bonuses received by teachers	a	None:7 Low:0 Moderate:0 High:0	
Likelihood of school closure	a	None:7 Low:0 Moderate:1 High:0	
Publication of results (Yes/No)[14]	a	9/10	
Publication of tables that compare school performance (Yes/No)	a	2/10	
Possible influence of national periodical assessments (lower secondary education) (2006) Source: Education at a Glance (OECD, 2008)[3] None/Low/Moderate/High[17]			
Performance feedback to the school	High	None:4 Low:1 Moderate:2 High:3	
Performance appraisal of the school management	Low	None:6 Low:2 Moderate:1 High:0	
Performance appraisal of individual teachers	None	None:8 Low:1 Moderate:0 High:0	
The school budget	Low	None:8 Low:1 Moderate:0 High:0	
The provision of another financial reward or sanction	None	None:9 Low:0 Moderate:0 High:0	
The assistance provided to teachers to improve their teaching skills	Moderate	None:5 Low:1 Moderate:3 High:0	
Remuneration and bonuses received by teachers	None	None:9 Low:1 Moderate:0 High:0	
Likelihood of school closure	None	None:9 Low:0 Moderate:0 High:1	
Publication of results (Yes/No)[14]	Yes	7/12	
Publication of tables that compare school performance (Yes/No)	No	2/12	
Use of achievement data for accountability (2009) (15-year-olds) Source: PISA Compendium for the school questionnaire (OECD, 2010b)[3]			
% of students in schools where the principal reported that achievement data is used in the following procedures			
Posted publicly	46.6	36.4	11/33
Used in evaluation of the principal's performance	42.7	35.5	13/33
Used in evaluation of teachers' performance	41.3	44.2	16/33
Used in decisions about instructional resource allocation to the school	61.4	32.2	5/33
Tracked over time by an administrative authority	81.0	65.2	10/33

SCHOOL EVALUATION			
Requirements for school evaluations by an inspectorate (lower secondary education) (2006) Source: Education at a Glance (OECD, 2008)[3] None/1 per 3+ years/1 per 3 years/1 per 2 years/1 per year/1+ per year	1 per 3 years	None:4 1 per 3 years:6 1 per year:1	1 per 3+ years:5 1 per 2 years:0 1+ per year:1
Possible influence of school evaluation by an inspectorate (lower secondary education) (2006) Source: Education at a Glance (OECD, 2008)[3] None/Low/Moderate/High[17]			
Influence on performance feedback			
Performance feedback to the school	High	None:0 Low:1 Moderate:1 High:10	
Performance appraisal of the school management	Moderate	None:0 Low:2 Moderate:3 High:7	
Performance appraisal of individual teachers	Moderate	None:1 Low:5 Moderate:2 High:3	
Financial and other implications			
The school budget	Moderate	None:5 Low:2 Moderate:2 High:1	
The provision of another financial reward or sanction	a	None:4 Low:4 Moderate:0 High:1	
The assistance provided to teachers to improve their teaching skills	High	None:1 Low:2 Moderate:6 High:2	
Remuneration and bonuses received by teachers	a	None:6 Low:1 Moderate:2 High:0	
Likelihood of school closure	a	None:2 Low:3 Moderate:2 High:2	
Publication of results (Yes/No)[14]	No	11/13	
Publication of tables that compare school performance (Yes/No)	No	1/12	

	Australia	Country Average[1]	Australia's Rank[2]
Requirements for school self-evaluations (lower secondary education) (2006) Source: Education at a Glance (OECD, 2008)[3]		None:6	1 per 3+ years:1
None/1 per 3+ years/1 per 3 years/1 per 2 years/1 per year/1+ per year	1 per year	1 per 3 years:1 1 per year:8	1 per 2 years:0 1+ per year:3
Possible influence of school self-evaluations (lower secondary education) (2006) Source: Education at a Glance (OECD, 2008)[3]			
None/Low/Moderate/High[17]			
Influence on performance feedback			
Performance feedback to the school	High	None:1 Low:2 Moderate:1 High:8	
Performance appraisal of the school management	Moderate	None:2 Low:2 Moderate:4 High:4	
Performance appraisal of individual teachers	Moderate	None:4 Low:4 Moderate:2 High:2	
Financial and other implications			
The school budget	Moderate	None:5 Low:2 Moderate:2 High:1	
The provision of another financial reward or sanction	a	None:4 Low:4 Moderate:1 High:0	
The assistance provided to teachers to improve their teaching skills	High	None:3 Low:2 Moderate:1 High:5	
Remuneration and bonuses received by teachers	a	None:5 Low:3 Moderate:0 High:1	
Likelihood of school closure	a	None:8 Low:0 Moderate:1 High:0	
Publication of results (Yes/No)[14]	No	4/14	
Publication of tables that compare school performance (Yes/No)	No	1/14	
Frequency and type of school evaluations (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers working in schools where school evaluations were conducted with the following frequency over the last five years			
Frequency of school self-evaluation			
Never	6.8	20.2	18/23
Once	25.3	16.2	5/23
2-4 times	14.1	18.3	16/23
Once per year	50.0	34.9	4/23
More than once per year	3.7	10.3	17/23
Frequency of external evaluation			
Never	21.2	30.4	14/23
Once	36.2	30.8	8/23
2-4 times	29.7	20.5	4/23
Once per year	10.7	11.4	11/23
More than once per year	2.2	7.0	=13/23
No school evaluation from any source	5.0	13.8	=16/23
Criteria of school evaluations (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers whose school principal reported that the following criteria were considered with high or moderate importance in school self-evaluations or external evaluations			
Student test scores	86.9	76.2	5/23
Retention and pass rates of students	81.9	70.8	8/23
Other student learning outcomes	94.8	78.9	1/23
Student feedback on the teaching they receive	69.0	72.7	14/23
Feedback from parents	88.3	77.3	5/23
How well teachers work with the principal and their colleagues	79.5	83.7	18/23
Direct appraisal of classroom teaching	58.8	71.1	19/23
Innovative teaching practices	78.6	76.7	=12/23
Relations between teachers and students	89.7	87.1	10/23
Professional development undertaken by teachers	87.3	81.5	7/23
Teachers' classroom management	79.6	80.7	14/23
Teachers' knowledge and understanding of their main subject field(s)	76.5	78.2	16/23
Teachers' knowledge and understanding of instructional practices in their main subject field(s)	70.8	77.5	18/23
Teaching of students with special learning needs	79.8	77.2	12/23
Student discipline and behaviour	88.0	83.6	7/23
Teaching in a multicultural setting	41.9	52.9	16/23
Extra-curricular activities with students (*e.g.* school plays and performances, sporting activities)	77.0	74.5	14/23
Impacts of school evaluations upon schools (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers whose school principal reported that school evaluations (external or self-evaluation) had a high or moderate level of influence on the following			
Level of school budget or its distribution within schools	76.4	38.0	2/23
Performance feedback to the school	96.2	81.3	2/23
Performance appraisal of the school management	88.5	78.7	6/23
Performance appraisal of teachers	64.9	71.1	16/23
Assistance provided to teachers to improve their teaching	86.8	70.3	3/23
Teachers' remuneration and bonuses	5.1	26.1	19/23

	Australia	Country Average[1]	Australia's Rank[2]
Publication of school evaluations (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers in schools where school evaluation results were :			
Published; or	75.7	55.3	4/23
Used in school performance tables	23.3	28.7	16/23
Accountability to parents (2009) (15-year-olds) Source: PISA Compendium for the school questionnaire (OECD, 2010b)[3]			
% of students in schools where principals reported that their school provides parents with information on:			
This child's academic performance relative to other students in the school	69.5	46.1	4/32
This child's academic performance relative to national or regional benchmarks	46.4	46.8	16/33
This child's academic performance of students as a group relative to students in the same grade in other schools	21.3	23.1	15/33
TEACHER APPRAISAL			
Frequency and source of teacher appraisal and feedback (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers who reported having received appraisal and/or feedback on their work with the following frequency from the following sources			
Feedback received from the principal			
Never	30.1	22.0	5/23
Less than once every two years	14.6	9.2	4/23
Once every two years	5.4	4.5	=7/23
Once per year	19.1	22.8	16/23
Twice per year	9.0	12.3	17/23
3 or more times per year	13.3	17.1	16/23
Monthly	3.8	6.6	=17/23
More than once per month	4.7	5.4	12/23
Feedback received from other teachers or members of the school management team			
Never	14.8	28.6	18/23
Less than once every two years	11.5	6.9	2/23
Once every two years	3.9	2.6	4/23
Once per year	16.9	13.3	6/23
Twice per year	10.7	9.7	=7/23
3 or more times per year	20.4	19.3	9/23
Monthly	10.8	10.4	10/23
More than once per month	10.9	9.1	=8/23
Feedback received from an external individual or body (*e.g.* external inspector)			
Never	73.8	50.7	4/23
Less than once every two years	12.3	19.0	13/23
Once every two years	3.0	5.4	16/23
Once per year	5.4	13.2	19/23
Twice per year	2.1	5.4	=16/23
3 or more times per year	2.2	4.3	=15/23
Monthly	0.6	1.2	=13/23
More than once per month	0.6	0.8	=8/23
Criteria for teacher appraisal and feedback (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers who reported that the following criteria were considered with high or moderate importance in the appraisal and/or feedback they received			
Student test scores	51.4	65.0	19/23
Retention and pass rates of students	51.8	56.2	15/23
Other student learning outcomes	62.1	68.4	16/23
Student feedback on the teaching they receive	58.4	72.8	22/23
Feedback from parents	54.7	69.1	22/23
How well they work with the principal and their colleagues	69.7	77.5	21/23
Direct appraisal of classroom teaching	59.9	73.5	19/23
Innovative teaching practices	66.5	70.7	18/23
Relations with students	80.1	85.2	19/23
Professional development undertaken	48.8	64.5	20/23
Classroom management	69.8	79.7	20/23
Knowledge and understanding of their main subject field(s)	72.4	80.0	18/23
Knowledge and understanding of instructional practices in their main subject field(s)	66.7	78.2	19/23
Teaching of students with special learning needs	41.2	57.2	22/23
Student discipline and behaviour	63.1	78.2	22/23
Teaching in a multicultural setting	29.1	45.0	19/23
Extra-curricular activities with students (*e.g.* school performances, sporting activities)	51.7	62.3	19/23

	Australia	Country Average[1]	Australia's Rank[2]
Outcomes of teacher appraisal and feedback (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers who reported that the appraisal and/or feedback they received let to a modest or large change in the following aspects of their work and careers			
A change in salary	5.6	9.1	13/23
A financial bonus or another kind of monetary reward	1.6	11.1	=18/23
A change in the likelihood of career advancement	16.9	16.2	7/23
Public recognition from the principal and/or their colleagues	24.1	36.4	20/23
Opportunities for professional development activities	16.7	23.7	16/23
Changes in work responsibilities that make the job more attractive	17.4	26.7	16/23
A role in school development initiatives (*e.g.* curriculum development group)	24.1	29.6	15/23
Actions undertaken following the identification of a weakness in a teacher appraisal (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers whose school principal reported that the following occurs if an appraisal of teachers' work identifies a specific weakness			
The principal ensures that the outcome is reported to the teacher			
Never	0.0	2.6	=19/23
Sometimes	3.1	9.5	16/23
Most of the time	21.6	25.8	16/23
Always	75.2	62.1	8/23
The principal ensures that measures to remedy the weakness in their teaching are discussed with the teacher			
Never	0.0	1.0	=11/23
Sometimes	4.0	9.4	14/23
Most of the time	30.4	30.7	9/23
Always	65.6	58.9	9/23
The principal, or others in the school, establishes a development or training plan for the teacher to address the weakness in their teaching			
Never	0.0	10.5	23/23
Sometimes	7.1	33.0	23/23
Most of the time	35.5	35.9	11/23
Always	57.5	20.6	1/23
The principal, or others in the school, imposes material sanctions on the teacher (*e.g.* reduced annual increases in pay)			
Never	91.9	86.0	=12/23
Sometimes	4.4	11.3	14/23
Most of the time	2.0	1.8	7/23
Always	1.7	0.9	6/23
The principal, or others in the school, report the underperformance to another body to take action (*e.g.* governing board, local authority, school inspector)			
Never	31.1	51.0	19/23
Sometimes	52.7	37.3	3/23
Most of the time	5.2	6.8	12/23
Always	11.0	4.9	3/23
The principal ensures that the teacher has more frequent appraisals of their work			
Never	3.2	9.0	17/23
Sometimes	38.9	34.5	9/23
Most of the time	37.0	41.3	14/23
Always	20.8	15.2	6/23
Teacher perceptions of the appraisal and/or feedback they received (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers who reported the following about the appraisal and/or feedback they had received in their school			
Appraisal and/or feedback contained a judgement about the quality of the teacher's work	68.1	74.7	18/23
Appraisal and/or feedback contained suggestions for improving certain aspects of teacher's work	55.4	58.0	18/23
Appraisal and/or feedback was a fair assessment of their work as a teacher in this school			
Strongly disagree	4.4	4.4	8/23
Disagree	10.1	12.4	13/23
Agree	66.7	63.3	=9/23
Strongly agree	18.8	19.9	12/23
Appraisal and/or feedback was helpful in the development of their work as teachers in this school			
Strongly disagree	6.2	5.6	8/23
Disagree	18.8	15.9	7/23
Agree	60.0	61.8	15/23
Strongly agree	14.9	16.8	13/23

	Australia	Country Average[1]	Australia's Rank[2]
Teacher perceptions of the personal impact of teacher appraisal and feedback (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers who reported the following changes following the appraisal and/or feedback they received in their school			
the following personal impact from appraisal and feedback			
Change in their job satisfaction			
A large decrease	3.3	2.5	=5/23
A small decrease	6.3	4.8	=4/23
No change	48.1	41.2	6/23
A small increase	34.2	37.3	17/23
A large increase	8.3	14.2	20/23
Change in their job security			
A large decrease	1.4	1.5	13/23
A small decrease	2.3	3.0	=14/23
No change	76.3	61.9	6/23
A small increase	12.7	21.8	19/23
A large increase	7.4	11.8	15/23
Impact of teacher appraisal and feedback upon teaching (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers who reported that the appraisal and/or feedback they received directly led to or involved moderate or large changes in the following			
Classroom management practices	24.1	37.6	18/23
Knowledge or understanding of the teacher's main subject field(s)	19.4	33.9	17/23
Knowledge or understanding of instructional practices	22.1	37.5	18/23
A development or training plan for teachers to improve their teaching	18.4	37.4	20/23
Teaching of students with special learning needs	14.2	27.2	22/23
Student discipline and behaviour problems	21.0	37.2	20/23
Teaching of students in a multicultural setting	8.1	21.5	21/23
Emphasis placed by teachers on improving student test scores in their teaching	24.7	41.2	19/23
Teacher appraisal and feedback and school development (lower secondary education) (2007-08) Source: TALIS (OECD, 2009b)[11]			
% of teachers who agree or strongly agree with the following statements about aspects of appraisal and/or feedback in their school			
In this school, the school principal takes steps to alter the monetary reward of the persistently underperforming teacher	7.1	23.1	20/23
In this school, the sustained poor performance of a teacher would be tolerated by the rest of the staff	42.8	33.8	5/23
In this school, teachers will be dismissed because of sustained poor performance	29.2	27.9	11/23
In this school, the principal uses effective methods to determine whether teachers are performing well or badly	48.7	55.4	15/23
In this school, a development or training plan is established for teachers to improve their work	54.5	59.7	13/23
In this school, the most effective teachers receive the greatest monetary or non-monetary rewards	9.2	26.2	20/23
In this school, if I improve the quality of my teaching I will receive increased monetary or non-monetary rewards	8.2	25.8	20/23
In this school, if I am more innovative in my teaching I will receive increased monetary or non-monetary rewards	9.0	26.0	=20/23
In this school, the review of teacher's work is largely done to fulfil administrative requirements	63.4	44.3	1/23
In this school, the review of teacher's work has little impact upon the way teachers teach in the classroom	61.4	49.8	3/23
Methods used to monitor the practice of teachers (2009) (15-year-olds) Source: PISA Compendium for the school questionnaire (OECD, 2010b)[3]			
% of students in schools where the principal reported that the following methods have been used the previous year to monitor the practice of teachers at their school			
Tests of assessments of student achievement	58.2	58.3	19/34
Teacher peer review (of lesson plans, assessment instruments, lessons)	65.2	56.3	14/34
Principal or senior staff observations of lessons	62.3	68.3	25/34
Observation of classes by inspectors or other persons external to the school	5.9	28.0	29/34

	Australia	Country Average[1]	Australia's Rank[2]
STUDENT ASSESSMENT			
Completion requirements for upper secondary programmes Source: Education at a Glance (2009)[3, 4, 12] ● Final examination /■ Series of examinations during programme / Δ Specified number of course hours and examination / ♦ Specified number of course hours only			
ISCED 3A	● (in some states) ■ Δ (in all states)	●21 ■19 Δ19 ♦3	
ISCED 3B	■	●6 ■8 Δ7 ♦0	
ISCED 3C	■	●17 ■18 Δ17 ♦1	
Student grouping by ability (2009) (15-year-olds) Source: PISA Compendium for the school questionnaire (OECD, 2010b)[3] % of students in schools where principals reported the following on student grouping by ability			
Student are grouped by ability into different classes			
For all subjects	2.3	9.4	25/33
For some subjects	89.8	37.4	=2/33
Not for any subject	7.9	50.4	29/33
Student are grouped by ability within their classes			
For all subjects	1.4	4.5	26/33
For some subjects	62.7	46.4	6/33
Not for any subject	35.9	47.0	26/33
Groups of influence on assessment practices (2009) (15-year-olds) Source: PISA Compendium for the school questionnaire (OECD, 2010b)[3] % of students in schools where the principal reported the following groups exert a direct influence on decision making about assessment practices			
Regional or national education authorities (*e.g.* inspectorates)	81.6	56.6	4/33
The school's governing board	6.9	29.6	31/33
Parent groups	10.0	17.3	=10/33
Teacher groups (*e.g.* staff association, curriculum committees, trade union)	42.8	58.1	22/33
Student groups (*e.g.* student association, youth organisation	19.1	23.4	=15/33
External examination boards	54.7	45.2	9/31
Responsibility for student assessment policies (2009) (15-year-olds) Source: PISA Compendium for the school questionnaire (OECD, 2010b)[3] % of students in schools where the principal reported the following groups have considerable responsibility in establishing student assessment policies			
Establishing student assessment policies			
Principals	82.2	63.5	10/33
Teachers	81.7	69.0	8/33
School governing board	12.6	26.5	24/33
Regional or local education authority	33.8	15.5	6/32
National education authority	4.5	24.3	29/33
Frequency of student assessment by method (2009) (15-year-olds) Source: PISA Compendium for the school questionnaire (OECD, 2010b)[3] % of students in schools where the principal reported the student assessment methods below are used with the indicated frequency			
Standardised tests			
Never	29.9	23.7	11/33
1-2 times a year	55.1	51.0	16/33
3-5 times a year	13.4	16.5	=18/33
Monthly	1.5	4.3	19/33
More than once a month	0.2	3.4	26/33
Teacher-developed tests			
Never	1.2	2.7	7/33
1-2 times a year	7.8	6.7	12/33
3-5 times a year	33.8	30.0	14/33
Monthly	32.0	27.6	13/33
More than once a month	25.2	33.3	20/33
Teachers' judgmental ratings			
Never	5.3	6.6	10/33
1-2 times a year	20.4	12.0	6/33
3-5 times a year	28.0	22.9	12/33
Monthly	15.2	15.7	15/33
More than once a month	31.1	42.2	21/33

	Australia	Country Average[1]	Australia's Rank[2]
Student portfolios			
Never	17.7	24.1	18/33
1-2 times a year	44.6	34.4	8/33
3-5 times a year	21.9	20.6	14/33
Monthly	9.0	10.4	15/33
More than once a month	6.8	9.3	11/33
Student assignments/projects/homework			
Never	1.0	1.5	=9/33
1-2 times a year	2.9	12.2	27/33
3-5 times a year	20.4	16.1	10/33
Monthly	22.0	13.6	5/33
More than once a month	53.8	56.5	19/33
% of students reporting the following on the frequency of homework (2000) (15-year-olds) Source: PISA Student Compendium (Reading) (OECD, 2001)[3]			
Teachers grade homework			
Never	8.9	14.9	18/27
Sometimes	46.9	44.2	=14/27
Most of the time	30.4	24.5	=7/27
Always	12.6	13.9	13/27
Teachers make useful comments on homework			
Never	24.3	23.5	12/27
Sometimes	50.7	50.1	14/27
Most of the time	19.5	19.2	11/27
Always	4.3	4.9	17/27
Homework is counted as part of marking			
Never	10.2	13.7	12/27
Sometimes	35.9	33.3	10/27
Most of the time	29.3	25.7	9/27
Always	23.3	24.7	13/27
Use of student assessments (2009) (15-year-olds) Source: PISA Compendium for the school questionnaire (OECD, 2010b)[3]			
% students in schools where the principal reported that assessments of students are used for the following purposes			
To inform the parents about their child's progress	99.0	97.5	14/33
To make decisions about students' retention or promotion	67.9	77.1	27/33
To group students for instructional purposes	81.2	49.8	4/33
To compare the school to district or national performance	60.4	53.0	13/33
To monitor the school's progress from year to year	82.9	76.0	19/33
To make judgements about teachers' effectiveness	44.3	46.9	16/33
To identify aspects of instruction or the curriculum that could be improved	86.1	76.7	12/33
To compare the school with other schools	43.5	45.4	16/33
% of students repeating a grade in the previous school year according to reports by school principals in the following levels (2009) (15-year-olds) Source: PISA Compendium for school questionnaire (OECD 2010b)[4]			
ISCED2	0.2	3.2	=24/29
ISCED3	0.4	4.5	23/29

Sources:

OECD (2001), *PISA 2000 Student Compendium (Reading)*, OECD, *http://pisa2000.acer.edu.au/downloads.php/*.
OECD (2008), *Education at a Glance, OECD Indicators 2008*, OECD, Paris.
OECD (2009a), *Education at a Glance, OECD Indicators 2009*, OECD, Paris.
OECD (2009b), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, OECD, Paris.
OECD (2010a), *Education at a Glance, OECD Indicators 2010*, OECD, Paris.
OECD (2010b), *PISA 2009 Compendium for the school questionnaire*, OECD, *http://pisa2009.acer.edu.au/downloads.php.*
OECD (2010c), *PISA 2009 Results: What Students Know and Can Do, Volume I*, OECD, Paris.

Data explanation:

m	Data is not available
a	Data is not applicable because the category does not apply
~	Average is not comparable with other levels of education
=	At least one other country has the same rank

TALIS is the OECD's Teaching and Learning International Survey which was implemented for the first time in 2007-08. The data provided concerns 23 countries. The results derived from TALIS are based on self-reports from teachers and principals and therefore represent their opinions, perceptions, beliefs and their accounts of their activities. Further information is available at *www.oecd.org/edu/talis*.

PISA is the OECD's Programme for International Student Assessment, which was undertaken in 2000, 2003, 2006 and 2009. 15-year-old students worldwide are assessed on their literacy in reading, mathematics and science. The study included 27 OECD countries in 2000, 30 in 2003 and 2006, and 34 in 2009. Data used in this appendix can be found at *www.pisa.oecd.org*.

Notes

1. The country average is calculated as the simple average of all countries for which data are available.
2. "Australia's rank" indicates the position of Australia when countries are ranked in descending order from the highest to lowest value on the indicator concerned. For example, on the first indicator "population that has attained at least upper secondary education", for the age group 25-64, the rank 17/30 indicates that Australia recorded the 17th highest value of the 30 OECD countries that reported relevant data.
3. The column "country average" corresponds to an average across OECD countries.
4.

Terms used to describe levels of education

ISCED classification (and subcategories)

ISCED 1 - Primary education

Designed to provide a sound basic education in reading, writing and mathematics and a basic understanding of some other subjects. Entry age: between 5 and 7. Duration: 6 years

ISCED 2 - Lower secondary education

Completes provision of basic education, usually in a more subject-oriented way with more specialist teachers. Entry follows 6 years of primary education; duration is 3 years. In some countries, the end of this level marks the end of compulsory education.

ISCED 3 - Upper secondary education
Even stronger subject specialisation than at lower-secondary level, with teachers usually more qualified. Students typically expected to have completed 9 years of education or lower secondary schooling before entry and are generally around the age of 15 or 16.

ISCED 3A - Upper secondary education type A
Prepares students for university-level education at level 5A

ISCED 3B - Upper secondary education type B
For entry to vocationally oriented tertiary education at level 5B

ISECD 3C - Upper secondary education type C
Prepares students for workforce or for post-secondary non tertiary education

ISCED 4 - Post-secondary non-tertiary education
Programmes at this level may be regarded nationally as part of upper secondary or post-secondary education, but in terms of international comparison their status is less clear cut. Programme content may not be much more advanced than in upper secondary, and is certainly lower than at tertiary level. Entry typically requires completion of an upper secondary programme. Duration usually equivalent to between 6 months and 2 years of full-time study.

ISCED 5 - Tertiary education
ISCED 5 is the first stage of tertiary education (the second – ISCED 6 – involves advanced research). At level 5, it is often more useful to distinguish between two subcategories: 5A, which represent longer and more theoretical programmes; and 5B, where programmes are shorter and more practically oriented. Note, though, that as tertiary education differs greatly between countries, the demarcation between these two subcategories is not always clear cut.

ISCED 5A - Tertiary-type A
"Long-stream" programmes that are theory based and aimed at preparing students for further research or to give access to highly skilled professions, such as medicine or architecture. Entry preceded by 13 years of education, students typically required to have completed upper secondary or post-secondary non-tertiary education. Duration equivalent to at least 3 years of full-time study, but 4 is more usual.

ISCED 5B - Tertiary-type B
"Short-stream" programmes that are more practically oriented or focus on the skills needed for students to directly enter specific occupations. Entry preceded by 13 years of education; students may require mastery of specific subjects studied at levels 3B or 4A. Duration equivalent to at least 2 years of full-time study, but 3 is more usual.

5. The year of reference for Australia is 2007.

6. Public expenditure includes public subsidies to households for living costs (scholarships and grants to students/ households and students loans), which are not spent on educational institutions.

7. Expressed in equivalent US$ converted using purchasing power parities.

8. Expenditure on goods and services consumed within the current year which needs to be made recurrently to sustain the production of educational services – refers to current expenditure on schools and post-secondary non-tertiary educational institutions. The individual percentage may not sum to the total due to rounding.

9. Public and private institutions are included. Calculations are based on full-time equivalents. "Teaching staff" refers to professional personnel directly involved in teaching students.

10. For Australia the indicators only includes general programmes in upper secondary education.

11. The column "country average" corresponds to an average across TALIS countries.

12. The column "country average" indicates the number of countries/systems, in which a given criterion is used, for example, regarding the indicator "Decision on payments for teachers in public schools". In the row "Management responsibilities in addition to teaching duties", ●12 ■18 Δ7 indicates that this criterion is used to determine the base salary in 12 countries/systems, to determine an additional yearly payment in 18 countries/systems and to determine an additional incidental payment in 7 countries/systems.

13. By "national examination" we mean those tests which do have formal consequences for students.

14. The column "country average" indicates the number of countries for which the indicator applies. For example, for the indicator "mandatory national examination is required" 4/20 means, that 4 countries out of 29 for which data is available report that mandatory national examinations are require in their countries.

15. By "national assessment" we mean those tests which do not have formal consequences for students.

16. In Australia, assessments are administered at the state level.

17. These measures express the degree of influence on the indicator: None: No influence at all, Low: Low level of influence, Moderate: Moderate level of influence, High: High level of influence. The column "country average" indicates the number of countries/systems, in which one of the given criteria is used.

Source Guide

Participation of countries by source

	PISA (OECD, 2000)	Education at a Glance (OECD, 2008)	Education at a Glance (OECD, 2009a)	TALIS (OECD, 2009b)	Education at a Glance (OECD, 2010a)	PISA Compendium (OECD, 2010b) PISA Results 2009 (OECD, 2010c)
Australia	●	●	●	●	●	●
Austria	●	●	●	●	●	●
Belgium (Flemish Community)		●	●	●	●	
Belgium (French Community)	●	●	●		●	●
Belgium (German Community)						
Brazil				●		
Canada	●	●	●		●	●
Chile					●	●
Czech Republic	●	●	●		●	●
Denmark	●	●	●	●	●	●
Estonia				●		●
Finland	●	●	●		●	●
France	●	●	●		●	●
Germany	●	●	●		●	●
Greece	●	●	●		●	●
Hungary	●	●	●	●	●	●
Iceland	●	●	●	●	●	●
Ireland	●	●	●	●	●	●
Israel						●
Italy	●	●	●	●	●	●
Japan	●	●	●		●	●
Korea	●	●	●	●	●	●
Luxembourg	●	●	●		●	●
Malaysia				●		
Mexico	●	●	●	●	●	●
Netherlands		●	●		●	●
New Zealand	●	●	●		●	●
Norway	●	●	●	●	●	●
Poland	●	●	●	●	●	●
Portugal	●	●	●	●	●	●
Slovak Republic		●	●	●	●	●
Slovenia				●		●
Spain	●	●	●	●	●	●
Sweden	●	●	●		●	●
Switzerland	●	●	●		●	●
Turkey		●	●	●	●	●
UK - England						
UK - Wales	●	●	●		●	●
UK - Norther Ireland						
UK - Scotland						
United States	●	●	●		●	●

OECD Reviews of Evaluation and Assessment in Education

Australia

How can student assessment, teacher appraisal, school evaluation and system evaluation bring about real gains in performance across a country's school system? The country reports in this series provide, from an international perspective, an independent analysis of major issues facing the evaluation and assessment framework, current policy initiatives, and possible future approaches. This series forms part of the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.

Contents

Chapter 1. School education in Australia
Chapter 2. The evaluation and assessment framework
Chapter 3. Student assessment
Chapter 4. Teacher appraisal
Chapter 5. School evaluation
Chapter 6. Education system evaluation

www.oecd.org/edu/evaluationpolicy