

OECD Reviews of Evaluation and Assessment in Education

Australia

MAIN CONCLUSIONS

Paulo Santiago, Graham Donaldson,
Joan Herman and Claire Shewbridge



This summary of main conclusions is drawn from *OECD Reviews of Evaluation and Assessment in Education: Australia*. The full report is available on the project's website: www.oecd.org/edu/evaluationpolicy.

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.

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.

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.