

Chapter 2

Standards, Assessment and Reporting

While we can be pleased to be significantly ahead of the OECD average and many OECD countries on all measures, we ought also to accept the challenge to match those ahead of us. We should not need the fiction of a quality crisis to inspire us to do even better.¹

2.1 A lay person is often struck by the fact that students may pass through six or even more years at school and remain functionally illiterate. More commonly, students may complete the final two years of secondary school and emerge with a restricted vocabulary, and without a firm grasp of how to construct a complex sentence. There is ample anecdotal evidence that such people have managed to make it through to higher education.

2.2 In this chapter the committee looks at current assessment programs, international tests which spotlight Australia's position, and their implications, benchmark tests, the need for national consistency in standards for levels of achievement, and ways of reporting these levels so as to have agreed understandings of what they mean.

Are standards declining?

2.3 Submissions state that there is a general decline in academic standards. The proportion of Australian students achieving only minimal literacy and numeracy skills are cited as evidence of the decline. The proportion of Australian students achieving below those levels required for effective functioning in adult society are also cited as evidence. The relatively poor performance in Trends in International Mathematics and Science Study (TIMSS) results was said to be most worrisome.²

2.4 University academics are in a strong position to see fluctuations in standards over a period of time. One told the committee:

The fact that academic standards are falling at schools and the university sector generally is undeniable. This is best seen at the second level universities and the less academic schools. Top universities, like ANU, Sydney, Melbourne, etc, will see this to a lesser extent because the shrinking market of well-trained school students will hit them last.³

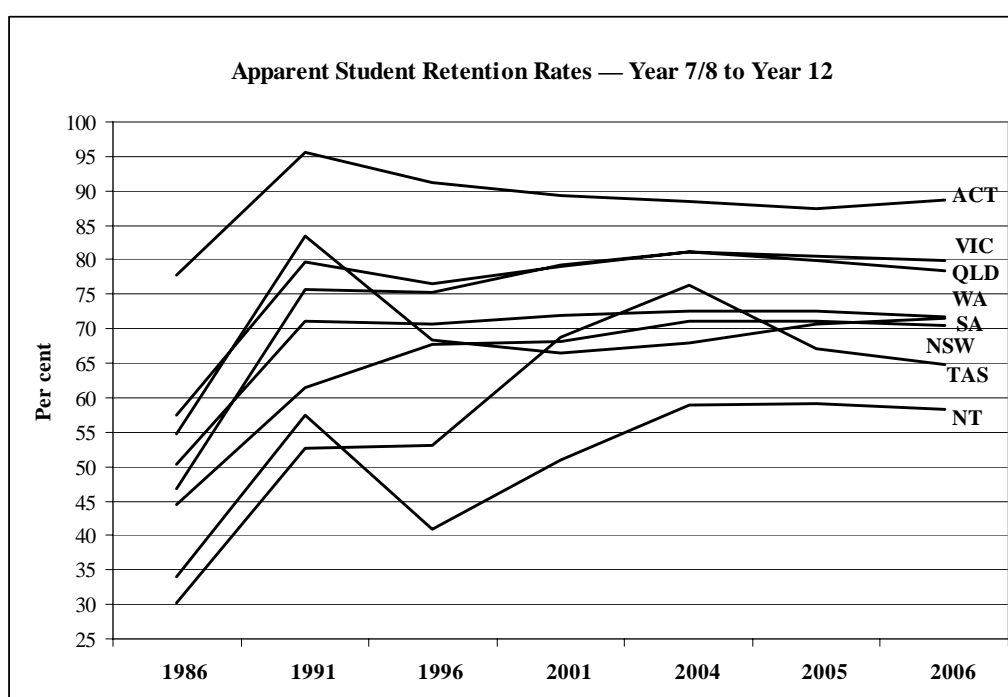
1 Barry McGaw, 'Resourced for a world of difference', *The Australian*, 1 August 2007, p. 25.

2 Australian Council for Educational Research, *Submission 38*, p. 2; Dr Kevin Donnelly, *Submission 9*, pp 3-4.

3 Professor Igor Bray, *Submission 6*, p. 1.

2.5 Another measure of the general decline in standards is in school completion rates. Australia has one of the world's lowest secondary school completion rates. This is behind East Asia, North America, Scandinavia, and much of continental Europe. Among 20-24 year olds, 17 per cent of Australians have neither completed secondary school nor are in education. For Norway, the corresponding figure is currently only 4 per cent.⁴

2.6 Some states and jurisdictions perform better than others in school completion rates and tertiary enrolments. For example, in Victoria, 85 per cent of 20-24 year olds had completed Year 12 or its equivalent in 2005, compared with 82.9 per cent in 1999. That was higher than the national average of 82.7 per cent. In 2006 the percentage of Year 12 school completers who enrolled in university increased from 46.1 per cent in 2003 to 47.4 per cent in 2007.⁵ A graph showing relative performance over recent years is set out below:



Source: Australian Bureau of Statistics, *Schools, Australia 2006*.

2.7 Other submissions argue strongly that claims of declining standards are irresponsible, and have branded it as political scaremongering serving only to undermine confidence in teachers and education systems across the country.

Scapegoating teachers undermines morale, excludes the very experience, deep understanding and insight that in situ experience brings, and presumes the solution without considering the problems.⁶

4 Australian Council for Educational Research, *Submission 38*, p. 2.

5 Dr Dahle Suggett, Department of Education, Victoria, *Committee Hansard*, Melbourne, 26 June 2007, p. 19.

6 Australian Education Union, *Submission 14*, p. 4.

2.8 As evidence of the lack of a general crisis, those of this opinion point to students' results in both national and international testing. The Australian Literacy Educators' Association denied that there is a problem with the teaching of literacy and instead argued that students just don't bother to learn literacy, or perhaps just don't bother to apply their literacy knowledge and skills.⁷

2.9 It makes more sense to isolate problem areas and deal with them appropriately. There are a number of quite distinct improvements that can be made to literacy and mathematics teaching. Some have to do with teaching method and with improvements to teacher training. Some have to do with curriculum and assessment.

National Assessment Programs

2.10 National assessment programs are intended to promote educational reform and enhance student outcomes. At present, there are three national assessment programs: science (samples of Year 6 students), civics and citizenship (samples of Year 6 & Year 10 students), and information and communications technology (ICT) literacy (samples of Year 6 & Year 10 students). These programs are conducted in a three-year cycle.

2.11 In 2003 the first sample assessment was conducted. The National Science Assessment determined that 58.2 per cent of students achieved at or bettered the 'proficient' standard, while 7.7 per cent of students achieved at higher proficiency levels.

2.12 In 2004 the second sample assessment was undertaken in Civics and Citizenship. Results from this assessment indicated that 50 per cent of Year 6 students achieved at or bettered the 'proficient' standard with 8 per cent performing at a higher proficiency. Among Year 10 students, only 39 per cent of students achieved at or bettered the 'proficient' standard and 5 per cent performed at a higher proficiency.

2.13 In 2005 the focus was upon ICT literacy. The results of this assessment are not yet available.

2.14 The national assessment programs do not comprehensively describe Australian students' levels of achievement in the three targeted areas. These programs apply only to a limited number of students, and the significance of their results depends upon a variety of contextual factors.

7 Dr Jan Turbill, Australian Literacy Educators' Association, *Committee Hansard*, Canberra, 11 July 2007, p. 13; Queensland Secondary Principals' Association, *Submission 56*, p. 1; Australian Education Union, *Submission 14*, p. 2.

English and Mathematics

2.15 Perhaps the best known and earliest programs of assessment were English and mathematics. These programs are more commonly known by reference to their assessment standards: the 'literacy and numeracy benchmarks'. The national benchmarks state the minimum acceptable standards of literacy and numeracy for Years 3, 5 and 7, and were approved by the Ministers of Education in 2000. Students in these years, and in some states and territories Year 9 students, participate annually in the English and mathematics national assessments. From 2008 the state-wide tests will be replaced by a national assessment program and include the Year 9 cohort.

2.16 The committee notes that the *2005 National Report on Schooling, National Benchmark Results, Reading Writing and Numeracy, Years 3, 5, and 7* is yet to be fully released. While the 2005 results, released in a preliminary paper, are detailed below, the 2004 results were utilised throughout the inquiry. The committee further notes that the results in 2004 and 2005 were consistent. Generally, student performance appears to be consistently high with a majority of students achieving at the benchmark level or higher in all states and territories. The trends in most areas tested show considerable stability over the life of the tests.

2.17 The benchmarking process is intended to support the National Goal that every child leaving primary school should be numerate and able to read, write and spell at an appropriate level. The development and implementation of the National Literacy and Numeracy Plan underpins this policy goal.

2.18 The literacy and numeracy benchmark tests seek to test the minimum standards of performance below which students will have difficulty progressing satisfactorily at school, and require increasing levels of proficiency from Year 3 though to Years 5 and 7.

2.19 The benchmark reporting builds an incremental picture of student achievement over time. Fundamentally, its purpose is to assist teachers' professional development and to enable interventionist support for students at risk.

	Year 3	Year 5	Year 7
Reading	92.7%	87.5%	89.8%
Writing	92.8%	93.3%	92.2%
Numeracy	94.1%	90.8%	81.8%

Source: MCEETYA, 2005 National Report on Schooling, National Benchmark Results, Preliminary Paper, Reading Writing and Numeracy, Years 3, 5, and 7

2.20 There were a few common trends throughout the 2005 results which bear mentioning. First, girls performed better than boys in reading and writing, whereas boys performed better than girls in numeracy. Secondly, the proportion of Indigenous

students achieving either at or above the benchmark level was substantially less than the proportion for non-Indigenous students. Thirdly, trend data suggests that Indigenous student performance is improving in literacy but not numeracy. While most students are reading, writing and spelling at an acceptable minimum level, there is room for improvement in some areas.⁸

2.21 The literacy and numeracy benchmark tests are of limited use as they do not apply to later stages of schooling. In fact, the results suggest that some students might complete compulsory schooling (Year 10) equipped with minimal literacy and numeracy skills. At present, there is no indication of what standards are actually achieved from Year 8 onward. It is conceivable that student achievement declines, particularly in the post-compulsory schooling years (Years 11–12) when curricula might be geared to matriculation requirements.

2.22 This lack of information will be partially remedied in 2007 with the anticipated endorsement and introduction of Year 9 benchmark standards and full cohort testing. The committee acknowledges MCEETYA's initiative in this regard, as well as its support for testing students' full range of abilities, rather than just the minimum benchmark standards.

National assessment program for literacy and numeracy

2.23 Notwithstanding the states' and territories' mixed commitment, they have raised concerns about financial, organisational and logistical costs which will be incurred with nationwide testing. For instance, Queensland has estimated that its costs in administering the assessment program will more than double. In Western Australia, Catholic and independent schools will receive no funding from the state to cover their costs of the testing.

International assessment programs

2.24 There are two internationally recognised assessment programs providing comparative achievement data across many countries. These were frequently referred to during the course of the inquiry. They test achievement in mathematics, reading, and science literacy: the Program for International Student Assessment (PISA), conducted every three years by the Organisation for Economic Co-operation and Development (OECD), which tests a sample of 15-year-old students, and the Trends in International Mathematics and Science Study (TIMSS), conducted every four years by the International Association for the Evaluation of Educational Achievement which tests a sample of students in Years 4 and 8.

PISA

2.25 PISA is a survey of the knowledge and skills of 15-year old students. In 2003, approximately 276 000 students in 41 countries participated in PISA which tested

8 Australian Association for the Teaching of English, *Submission 3*, p. 3.

mathematical, scientific and reading literacy, as well as an additional area, problem solving. PISA assesses students' ability to apply their knowledge and skills to real life problems and situations, rather than how well they have learned a specific curriculum.

2.26 Australia's PISA 2003 results were described as good to excellent in each of the tested areas. In mathematical literacy, four countries outperformed Australia, an increase of two countries following the PISA 2000 assessment. Three countries returned significantly higher results in scientific literacy compared with two countries in PISA 2000. In reading literacy, only one country achieved significantly higher results than Australia, a result identical to the results from PISA 2000. Problem solving was tested for the first time in 2003 and the results indicate that four countries outperformed Australia.

2.27 Generally, Australian students' results were consistently and significantly above the OECD average. The Australian Mathematical Sciences Institute submission noted that PISA results are frequently quoted as indicating that Australian students are performing well in mathematics compared with other nations. While this was commendable, it is not a valid assessment of the mathematics knowledge as only a fragment of mathematics' curriculum is tested. Some of the questions are effectively general aptitude tests rather than mathematical ones.

2.28 The results from PISA are often hailed as evidence of Australian students' high academic achievement in the areas of literacy and numeracy.⁹ While this appears to be true for students, the committee was constantly reminded in evidence about that proportion of students who did not perform so well in the PISA assessment.

TIMSS

2.29 TIMSS is different from PISA in that it is closely linked to the mathematics and science curricula of participating countries. According to the Australian Mathematical Sciences Institute, TIMSS is the best guide as to how Australia is comparing internationally in mathematics because it concentrates on content. It is designed to measure trends in students' knowledge and abilities.

2.30 In 2003, 46 countries participated in TIMSS with Australian students in fourth and eighth grade undertaking the assessments. By Year 8 the curriculum and expectations of students are similar internationally, and differences in school starting ages have had time to even out. In addition, the Year 8 TIMSS tends to have more countries involved. Many educationists regard this test as providing much more useful information than PISA. Some countries, eg highly performing ones such as Singapore, participate in TIMSS but not in PISA.¹⁰ The committee notes that this is probably the

9 For instance, Australian Association for the Teaching of English, *Submission 3*, p. 2; Association of Principals of Catholic Secondary Schools in Australia, *Submission 16*, p. 2.

10 Ms Yvonne Meyer, *Submission 17*, p. 2; Australian Mathematical Sciences Institute, *Submission 42*, p. 2.

reason why PISA results are generally more favourably perceived than TIMSS, which gives rise to as much concern as it does gratification.

2.31 Australian TIMSS results show that there is much to be concerned about. Two points stand out: the first is the long tail of under-achievement indicating a high percentage of students who, early in their secondary education, are unlikely to have acquired the necessary background skills for intermediate and advanced level mathematics courses at Years 11 and 12; the second is the low percentage in the highest level compared with the leading countries, bearing out the view of senior teachers and academics that expectations of Australian students are mostly ‘average’ and that they are insufficiently motivated and challenged.¹¹

2.32 Australia's 2003 TIMSS results showed that fourth-grade students performed above the international average in both science and mathematics. However, the average score in mathematics was not significantly higher than the international average. In both tested areas there was negligible improvement over an eight year period. While Australia's results were similar to some industrialised countries, Australian students did not perform as well as students from the United States and Britain.

2.33 Eighth-grade students performed well above the international average in both science and mathematics. In science there was a reasonable improvement on the 1995 TIMSS results, whereas there was a slight decline in the average mathematics score. While the Australian results were generally comparable to some industrialised countries, they were arguably lower than the Asia-Pacific regional average.

General responses to the international test results

2.34 The committee was told the Australian model for the teaching of literacy is viewed favourably abroad, so much so that some countries which are improving in PISA are moving toward similar models.¹² The committee notes the majority of submissions and evidence affirmed and applauded the strong performance of most students in PISA and TIMSS. The majority of submissions and evidence, however, made a strong point in identifying the large tail of students, who are not meeting the minimum benchmarks.

30 per cent of Australian 15-year olds [are] not achieving a level of reading proficiency regarded by the OECD as being needed to meet the demands of lifelong learning in a rapidly changing knowledge-intensive society. Of even greater significance is that 11.8 per cent of 15-year-olds—that is about 30,000 students each year—achieve only at or below level 1 in these tests.¹³

11 Australian Mathematical Sciences Institute, *Submission 42*, p. 2.

12 Mr Mark Howie, Australian Association for the Teaching of English, *Committee Hansard*, Canberra, 11 July 2007, pp 19-20.

13 Mr Bill Burmester, DEST, *Committee Hansard*, Canberra, 11 July 2007, p. 26.

2.35 The committee is most concerned that these results are put in perspective. There appears to be a large proportion of students who are not achieving a minimal standard of literacy and numeracy and whose opportunities in life will be curtailed as a result of that failure. Despite protestations to the contrary, the committee fears that they may encourage complacency.

2.36 In identifying the source of the problem, Professor Bill Loudon from the University of Western Australia told the committee:

We do very well with the top third of the population...If there is a black hole it is in the bottom half of the population academically and year 12, and throughout for the bottom half of kids we just do not have it right anywhere beyond years 3 or 4...In terms of standards, kids in the bottom quartile of mathematics performance at year 5 probably learn no more mathematics, although they do another five years of mathematics. Kids who are in the top quartile in year 5 mathematics—in the top five per cent particularly—become marvellously facile in mathematics, continue to learn every year and then go off to university and do university mathematics. But there are a lot of kids who are just marking time. The economy has no place for them, schools are not really organised for them and do not find them easy to teach. So that is where the standards problems are.¹⁴

2.37 This observation was supported by Professor Greg Robson from Edith Cowan University:

The problem we have across schools and school systems is—to use a sporting analogy—that it is a patchy performance. It is not consistently high in as many places as it should be. We have pockets—and they are reasonably substantial pockets—of high performance accompanied by areas where we know we need to do much better.¹⁵

2.38 The Australian Education Union agreed:

The evidence, looked at rationally, overwhelming indicates that the major problem facing Australia is low achievement associated with students from low SES backgrounds, including, but not limited to, those from Indigenous backgrounds and those in rural and remote areas.¹⁶

2.39 In the Northern Territory achievement levels are consistently well below those of other states and territories. This is partially due to the high proportion of indigenous students and a widely dispersed population with many small communities. However, these problems exist to some degree within other jurisdictions, such as Queensland, Western Australia and New South Wales. The committee believes that the serious

14 Professor Bill Loudon, *Submission 73*, pp 1-2.

15 Professor Gregory Robson, Edith Cowan University, *Committee Hansard*, Perth, 2 July 2007, p. 37.

16 Australian Education Union, *Submission 14*, p. 4.

problems afflicting education in the Northern Territory are due also to school availability and notoriously poor attendance levels.

2.40 Socio-economic status does not appear to be a relevant factor in those countries which perform better than Australia in PISA and TIMSS. However, the Australian Council for Educational Research (ACER) indicated to the committee that the socio-economic background of students is not necessarily the determining factor of low achievement:

Increasing variability across the years of school sometimes is reflected in growing gaps between students from lower and higher socio-economic backgrounds and between Indigenous and non-Indigenous students. It is important to note that although students' socioeconomic background is correlated with school achievement, the correlation is not high (generally less than 0.3).¹⁷

2.41 The apparent problem of low socio-economic status has been resolved at the school level in some schools. For instance, in Victoria, Catholic school enrolments are very evenly distributed across income and social groups, being almost 10 per cent in each SES decile. Yet the academic results achieved by those schools are higher than might otherwise be expected. The committee believes that the socio-economic status factor is surmountable, as it has been in past generations which have seen an 'aspirational' cohort rise from their working class origins. The difficulty for schools and teachers is to motivate students to develop an interest in their own educational growth.¹⁸

2.42 Another instance of the significant variability in students' levels of achievement is the 7 per cent of Australian girls and 17 per cent of Australian boys who perform at the lowest international literacy standard. There is no obviously apparent reason for the gender disparity, but might simply be attributable to the disengagement of boys in classroom activity. In Year 8 mathematics only 7 per cent of Australian students perform at an advanced level compared with 44 per cent of Singaporean students. According to Professor Michael O'Neill, this evidences a perennial tension between process and content.¹⁹

We have this tension in teaching and in schooling where we have had less emphasis on core knowledge and the core disciplines and greater emphasis on applied knowledge and process.²⁰

17 Australian Council for Educational Research, *Submission 38*, p. 1.

18 Catholic Education Commission of Victoria, *Submission 15*, p. 2.

19 Australian Council for Educational Research, *Submission 38*, p. 1; Dr Phillip McKenzie, Australian Council for Educational Research, *Committee Hansard*, Melbourne, 25 June 2007, p. 45.

20 Professor Michael O'Neill, University of Notre Dame Australia, *Committee Hansard*, Perth, 2 July 2007, p. 35.

2.43 The committee understands this to mean that test results show that Australian students know less as a consequence of their pursuit of 'relevance'. While all mathematics experts talk about the need for 'deep knowledge and understanding' it appears that this can only come about through children undertaking tasks which would be criticised in this country as being 'mechanical', as if that disadvantaged them. It is an issue that will be taken up in a later chapter.

2.44 The rigour and validity of the PISA assessment was also called into question. In literacy, PISA does not mark students down for errors in spelling, grammar, punctuation and style. More importantly, in mathematics, PISA assesses life-skills rather than concepts, skills and preparation for further study.

2.45 Although Australian students performed well overall in TIMSS 2003, there is concern over the apparent lack of improvement in comparison to other countries. With the exception of Year 8 science, levels of performance of Australian students has been maintained but not improved. Other countries, by comparison, are doing better now than they were previously.²¹

Australia's economic competitors are outperforming us. This is a national concern as well as providing Australian students with an education that will place them in a weaker position in the global world in which they live and work.²²

Standards

2.46 The committee noted a number of submissions presenting arguments that the inquiry, like the prevailing school policies, was much too preoccupied with standards. Some of these views are set down and commented on below. The reference to the word 'standards' provoked adverse comment from some submitters. It was argued that the focus was misdirected, and that the associated testing regimes were contrary to excellence in teaching and that 'standards' are themselves a construct of convenience:

['Standards'] appear to be primarily constructs of convenience that express themselves mainly in statistical terms (eg benchmarks) and they reflect certain expectations of those who have a special interest in the capabilities of the graduates moving out of the respective stages of the schooling process (ie Yr 2, Yr 6, Yr 10, Yr 12)... The focal point in the debate is 'standards' but this disguises the core endeavour of effective educational practice: a disposition to apply the outcomes of one's learning to the multitude of real-life contexts that will punctuate one's life.²³

21 Queensland Department of Education, Training and the Arts, *Submission 54*, p. 27; Australian Education Union, *Submission 14*, p. 13; Professor Gregory Robson, Edith Cowan University, *Committee Hansard*, Perth, 2 July 2007, p. 37.

22 Dr John Ridd, *Submission 4*, pp 1-2.

23 Australian Association of Christian Schools Ltd, *Submission 34*, pp 2 & 6.

2.47 In supporting standards-based curricula the committee accepts that it has a special interest in the capabilities of those who progress successfully through the stages of their schooling. The future depends on this happening. There is no philosophical conflict between the goal of reaching desired levels of academic success and learning to cope with real life. The goals of schooling are necessarily wide.

The measure of a student's achievement and success is not simply a grade or a number. Standards of academic achievement are too often defined in a narrow, quantitative way. Standards should be clearly justified, defined and criterion-referenced and as a general rule, exist to support authentic and deep learning.²⁴

2.48 The committee would not argue that success must always be measured in academic terms. Individuals learn when they are ready. The committee's view is that standards should be justified, defined and criterion referenced. The problem is that many schools and systems have not yet reached this point. The committee would generally agree that the setting down of standards—what students are expected to know and understand in their various subjects—is important if we are to ensure that particular levels of competence are comparable across the country, and that they can be reported on accordingly. Standards ensure an acceptable minimum or average performance equating to competence. They are not set to ensure homogeneity. The committee accepts the views expressed by the Association of Heads of Independent Schools of Australia who submitted:

Data should be at the school, regional and national level and must be used to provide standards as reference points, not used for standardisation. Standardisation constrains the professional responses that schools or classroom teachers are able to provide. Standardisation is antithetical to excellence and it will not provide the skills of literacy numeracy and scientific knowledge, attitudes and behaviours that adults of the mid 21st century will require.²⁵

2.49 The committee also acknowledges the value of opinion expressed by the Queensland Catholic Education Commission, and others, who stressed that education was broader than exams:

Obviously test results have a small part to play in the overall educational scene...Education is about much more than just testing young people. If you get down to that notion of testing a very limited slice of the curriculum and putting great value in those results, excluding everything else, what you risk is cutting out the richness and the broadness of a young person's curriculum and cutting out some of their local context and how important that is. So, yes, test results have a part, but it is a part of a whole big picture

24 Lutheran Education Australia, *Submission 41*, p. 5.

25 Association of Heads of Independent Schools of Australia, *Submission 18*, p. 1.

that looks at the development of a young person socially, emotionally, physically and intellectually.²⁶

2.50 The committee is aware of the dangers of overassessment, as recent British experience has shown, just as it is aware that not all things learnt at school can or should be tested. But the committee also believes that some educators place too little emphasis on testing, on the basis of certain philosophical issues they have concerning competitiveness and freedom from anxiety. Both anxiety and competitiveness are life-skill challenges which should be encountered and dealt with in a friendly and supportive school environment.

2.51 Whatever the view taken of 'standards' the committee believes they serve a useful function in that they identify minimum performance targets. This allows for current levels of achievement to be identified and for learning to be customised to serve the needs of individual students. As the ACER repeatedly stresses, it is all about promoting growth. That is also the purpose of benchmarking tests:

When the [benchmarking] was introduced, it was introduced with a view to realising the data's potential for diagnosis and timely intervention and improvement, so it had a strong equity agenda. That requires that the shift of emphasis be less on measurement and more on using the data to inform classroom pedagogy and diagnosis of need.²⁷

2.52 The committee has been told that among educators there is a fundamental belief that all students are capable of progressing beyond their current levels of achievement. The challenge is to understand each student's current level of achievement and to provide opportunities likely to facilitate further growth. First and foremost, this requires sound and reliable information or data.

It is vital that teachers are provided with standards-based assessment instruments...constructed and calibrated on nationally consistent, common measurement scales that are qualitatively described.²⁸

Progressive failure

2.53 The long performance tail identified in international testing suggests that early in secondary school there is already a high percentage of students who are unlikely to have acquired the necessary foundation skills. Worse, the gap between students meeting the international benchmarks and those who do not, increases as students progress through school. In Western Australia, for example, the percentage of children meeting the literacy benchmark for Years 3, 5 and 7 are 92.8 per cent, 90.5 per cent

26 Mrs Diane Anderson, Queensland Catholic Education Commission, *Committee Hansard*, Brisbane, 5 June 2007, p. 65. Also, Professor Claire Wyatt-Smith, Griffith University, *Committee Hansard*, Brisbane, 5 June 2007, p. 90.

27 Professor Claire Wyatt-Smith, Griffith University, *Committee Hansard*, Brisbane, 5 June 2007, pp 85-86.

28 Australian Council for Educational Research, *Submission 38*, p. 4.

and 81 per cent: a declining average. This suggests that Australia is failing to properly address the problems of illiteracy in students.²⁹

Benchmark testing

2.54 Considerable concern has been expressed in both submissions and evidence about the validity of benchmark testing.

2.55 These tests are intended to test the minimum standards of performance below which students will have difficulty progressing satisfactorily at school. It is intended as a 'safety net' to identify students at risk of failure. As one experienced Queensland educator told the committee:

The whole purpose of a test is that they send a signal. The moment they send that signal there should be immediate allocation of appropriate resources to the areas where there are deficiencies...There is no point in having testing unless it is immediately followed by remedial measures...I do not think that happens to such a large extent.³⁰

2.56 It is argued in some circles that this focus on minimum achievement in basic areas can lead to teachers giving more attention to students around the threshold benchmark, rather than all students across a broader curriculum. The committee considers this to be a spurious objection, if only because it assumes a lack of professionalism on the part of teachers. Testing has an obvious remedial purpose in primary school years, and it is not a valid criticism that benchmark testing does not trigger remedial attention.

Criticism of benchmark testing

2.57 Some submissions criticised the standards of achievement indicated by the 'benchmarks'. Not everyone agrees that benchmark tests identify students at risk. As one parent submitted:

Each year the states and territories publish information proclaiming that almost all students 'meet the benchmark'. However, the 'benchmark' is an arbitrary illusion that can be manipulated in order to deliver whatever result is required for whatever purpose. To announce that most students 'meet the benchmark' is a meaningless statement that provides false assurances to the general public.³¹

2.58 This assertion was strenuously rejected by the Victorian Curriculum and Assessment Authority which helps to administer the tests:

29 Australian Mathematical Sciences Institute, *Submission 42*, p. 3; Ms Christine Cook, Department of Education and Training, Western Australia, *Committee Hansard*, Perth, 2 July 2007, pp 73-74.

30 Professor Kenneth Wiltshire, *Committee Hansard*, Brisbane, 5 June 2007, pp 14-15.

31 Ms Yvonne Meyer, *Submission 17*, p. 2.

At the moment in the national testing there is only one benchmark, and it is a minimum proficiency one. It is admittedly not at a spectacularly high level. The point of establishing a minimum proficiency is to give a warning sign, if you like, that if a student is below that then they genuinely need additional support. So typically we have seen figures in the reports showing that in the high 80s to 90 per cent of students at most levels reach the benchmark. They are very consistent figures around the country. They vary up and down by one or two per cent by and large, but they are reasonably consistent...There is certainly no manipulation of the data. They are objectively marked. They are subject to quality assurance processes. The data are published freely back to schools...It is a transparent process as far as schools are concerned...It is run according to standard international assessment processes and we use experts to do it.³²

2.59 Professor Claire Wyatt-Smith from Griffith University was similarly critical of the minimal benchmark standards:

Teachers have indeed gone away from using identification of students at the thresholds on literacy coming from the test because they see they are so low that students who are above the minimum are at educational risk in their schools. I suggest that there is a need to look for what the minimum really represents now.³³

2.60 The education unions submitted that national benchmark tests are often used to place responsibility on teachers for 'poor' outcomes. It was argued by the Independent Education Union that such testing does not respect or involve the expertise and professional judgement of the teaching profession, nor does it have teachers' full support and confidence.³⁴ There was some confirmation of this from education faculty academics from Griffith University:

The data is not routinely used by teachers in conjunction with their own classroom assessment evidence. This is largely a result of the teachers' lack of professional development about how they might use the data for improvement (as distinct from measurement) purposes. In effect, the reported data are seen as a series of terminal points instead of a means of tracking performance for individuals and groups over time. The data is therefore being used for neither its intended purpose, nor to generate informed debate...There is also research evidence showing that quality literacy and numeracy assessment by teachers can lead to improvement for all students. There is no doubt that socioeconomic disadvantage is a key consideration in analysing student achievement data. However, this does not sufficiently explain continued or prolonged underperformance in certain

32 Mr John Firth, Victorian Curriculum and Assessment Authority, *Committee Hansard*, Melbourne, 26 June 2007, p. 20.

33 Professor Claire Wyatt-Smith, Griffith University, *Committee Hansard*, Brisbane, 5 June 2007, pp 86-87. Also, Dr Kerry Hempenstall, *Committee Hansard*, Melbourne, 25 June 2007, p. 21.

34 Independent Education Union of Australia, *Submission 55*, pp 8-9.

geographic areas and groups in our society; poverty does not equate to inevitable underperformance.³⁵

2.61 On the face of it, the committee rejects these criticisms. Self interest dictates these criticisms. It was suggested that if the data were more 'user friendly' and teachers were properly trained in its use, it might be better used. This is a priority task for system and school administrators. It occurs to the committee that it is very surprising that schools would endure the likely disruption of school routine to administer these tests and then not bother to use the results. The committee heard no comment from school principals on this issue. It notes confirmation in Griffith University's submission from the dean of the faculty at the Brisbane hearings:

The improvement data nexus was not followed through to the hands of teachers where that could be realised, and in fact teachers were the recipients of the information rather than the users of it. They became accountability measures rather than pedagogical devices.³⁶

2.62 The committee noted that teachers tended to regard mandatory testing as extraneous:

Any primary schoolteacher worth their salt can look around the class of 28 and say: that kid needs this; that kid needs that. They do not need a test to all them that. What they need is the resources to help those youngsters through.³⁷

2.63 The Australian Literacy Educators' Association pointed that within the classroom the teacher is constantly assessing a student to determine whether a particular strategy is working.³⁸ The committee acknowledges that benchmarking policy probably has, at its core, an element of supervision. It is a case of keeping teachers up to the mark. No government or school system, however, would be likely to put it in those terms.

Limitations of standardised tests

2.64 Another primary concern expressed in submissions was that standardised testing is limited. The Australian Primary Principals' Association noted that the use of multiple-choice questions was a limited mechanism which signalled an indifference to the role of the curriculum. The testing methods meant that much of the syllabus that

35 Dr Glenn Finger et al, *Submission 46*, p. 6.

36 Professor Claire Wyatt-Smith, Griffith University, *Committee Hansard*, Brisbane, 5 June 2007, p. 86. Also, Ms Janine McIntosh, International Centre of Excellence for Education in Mathematics, *Committee Hansard*, Melbourne, 26 June 2007, p. 34.

37 Mr Ian Ferguson, Queensland Secondary Principals' Association, *Committee Hansard*, Brisbane, 5 June 2007, p. 35. Also, Dr Ruth Fielding-Barnsley, *Committee Hansard*, Brisbane, 6 June 2007, p. 4; Dr Kerry Hemenstall, *Committee Hansard*, Melbourne, 25 June 2007, p. 21.

38 Dr Jan Turbill, Australian Literacy Educators' Association, *Committee Hansard*, Canberra, 11 July 2007, p. 14.

was really important to students, such as thinking mathematically and using language properly, could not be tested.³⁹ A similar point was made by the Australian Education Union, which submitted that much of what is important in schooling is not measured by standardised tests. The problem with them was that they focused attention on those areas of the curriculum that are tested, so that what is tested becomes what is viewed as important. Consequently, the range of things to be tested was expanded in order that they be seen as important.⁴⁰ The president of the Australian Education Union explained to the committee:

In a normal circumstance a teacher uses a test to tell the teacher about what the child is learning and to inform the teacher about future remediation. That is one of the problems with those standardised tests: they do not do that. By the time the results come back it is probably too late to do anything about that particular class. It provides a useful snapshot about where your class is in relation to the rest of the state or the rest of the country. It should not be used to do anything more than that...We believe that the bulk of the results could be achieved by sample testing rather than by testing the whole cohort.⁴¹

2.65 Another major criticism was that standardised testing could result in a culture of teachers teaching simply to pass the test.

If there are national tests, have no doubt our teachers will teach the test. They want the children to succeed. They want them to look good in the eyes of their peers. They want their school to have good data. So teachers will teach the test at the cost of professional freedom and at the cost of creativity in the classroom and so on.⁴²

2.66 The committee believes that system administrators and schools should review procedures in the light of classroom experience.

39 Dr Ian Chambers, Australian Primary Principals' Association, *Committee Hansard*, Sydney, 17 May 2007, p. 21.

40 Australian Education Union, *Submission 14*, p. 7; Professor Alan Reid, Australian Curriculum Studies Association, *Committee Hansard*, Melbourne, 25 July 2007, p. 2. For example, some qualities which are difficult to measure with standardised testing include creativity, critical thinking, resourcefulness, curiosity, spontaneity, et cetera.

41 Ms Pat Byrne, Australian Education Union, *Committee Hansard*, Melbourne, 25 June 2007, pp 7-8 & 9. Also, Mr Chris Watt, Independent Education Union of Australia, *Committee Hansard*, Melbourne, 26 June 2007, p. 12.

42 Mr Ian Ferguson, Queensland Secondary Principals' Association, *Committee Hansard*, Brisbane, 5 June 2007, p. 38. Also, Mr Chris Watt, Independent Education Union of Australia, *Committee Hansard*, Melbourne, 26 June 2007, p. 13; Australian Primary Principals' Association, *Submission 43*, p. 15 noting that 'teaching to the test' subverts the fundamental purpose of education by stifling genuine learning.

Benchmark testing – the committee's final word

2.67 Notwithstanding these comments, formed by knowledge and experience, the committee believes that some form of standardised diagnostic testing is necessary in all schools. It agrees with the Australian Primary Principals' Association that care needs to be taken that testing and assessment remain firmly linked to the purpose of achieving improvements in learning for students. Nor should the measurement of outcomes be an end in itself, as distinct from a means to achieve continuing improvements for students.⁴³ The committee accepts that refinements should be made, and that these should follow a process of consultation with teachers which appears to have so far been neglected. It finds the indifference of teachers to the testing regime—and we don't really know the extent of this—to be significant because it emphasises a point made elsewhere in this report to the effect that teachers can be led but they cannot be driven. Benchmark testing has a place in a national curriculum, but it should be part of a negotiated whole-of-curriculum approach.

'League tables'

2.68 Under the budget measures announced for 2007-08, the Government has announced that in the next quadrennium schools will have to report on their performance in literacy and numeracy benchmark tests.

2.69 Some witnesses expressed support for publishing lists of schools in rank order of academic performance, whereas others were emphatically opposed to the idea. It appears to be contrary to the spirit of the times. Many years have passed since the rank order of students in the NSW Leaving Certificate were published in the newspapers, including separate lists of those ranked in subjects at honours level, together with all successful students and their grades, identified with the schools they attended.

2.70 Schools appear nervous about having their students' assessed standards identified because of the concept of 'league tables'. The objection was that the data could be used to make unfair comparisons of schools. A number of variables affect the quality of education and schools indicated as 'underperforming' might be adversely affected by factors beyond their control.⁴⁴ This sensitivity appears to be directly targeted by the Government's policy, agreed to by COAG, to identify schools with the achievement levels of their students.

2.71 Most teaching bodies appearing before the committee expressed the view that such publication was unfair.

If you are in the top 10, that is fantastic but if you are a bit below that, that is whatever it is. I do not know how we get across to our parent body or to anyone else who might pick up the paper and have a look at where my

43 Australian Primary Principals' Association, *Submission 43*, p. 11.

44 Professor Max Coltheart, *Committee Hansard*, Sydney, 17 May 2007, p. 5; Australian Education Union, *Submission 14*, p. 33.

school sits that I had a year 8 student who when he came into my school could not read but still passed his year 12 English. How do we measure and report on that? I think that is a greater achievement perhaps than getting all your kids past year 12 in the end.⁴⁵

2.72 Interestingly, this viewpoint seems to be most strongly expressed by Catholic systems and by representatives of Lutheran and evangelical Protestant schools, many of which are newly established and sometimes struggle to find experienced teachers.

2.73 Despite these comments, the committee sees some public benefit in parents and the wider community being able to rank and compare schools against each other in some key areas of comparison, for instance academic achievement. This would allow parents to have a more informed choice in deciding which school is best for their child. It would also apply healthy competitive pressure to improve their relative rankings.

Reporting progress

2.74 The committee acknowledges that there are wide variations in students' levels of achievement. Children begin school with different levels of individual development and school readiness. They also learn at different rates, with some students requiring more time to learn than their peers. The gaps in levels of achievement widens over time so that, for instance, by Year 5 the top 10 per cent of children in reading are at least five years ahead of the bottom 10 per cent of readers.⁴⁶

2.75 The variation in students' skills levels upon transition from primary school to secondary school can be highly evident. As with universities and matriculating students, teachers are sometimes compelled to re-teach skills.

2.76 It is essential that students have a firm grasp on the fundamentals, without which it is impossible to build further knowledge, skills and understandings. A failure to grasp the basics can be a fatal flaw in education, and limit the range of options and opportunities for further success in life. Yet the word 'failure', is taboo in education circles, as one academic explained:

We have almost expunged the word 'failure' from our vocabulary in this country and in others in education. I think it is time we used the 'f' word again...In the interests of self-esteem we belittle success. We have demeaned success because we have expunged failure. Success is valued only at the risk of failure.⁴⁷

45 Mr Mark Rathjen, Living Waters Lutheran College, *Committee Hansard*, Perth, 2 July 2007, pp 55-56.

46 Australian Council for Educational Research, *Submission 38*, p. 1.

47 Professor Michael O'Neill, University of Notre Dame Australia, *Committee Hansard*, Perth, 2 July 2007, p. 44.

2.77 An experienced former teacher also expressed misgivings about the tendency of schools to protect the self-esteem of students:

Too often, we do not let them fail, take risks or become creative because we are so busy with following very clear guidelines, protecting them and so forth. What we are losing here is the ability of students to take care of themselves. I think that will have a very big impact on us as well.⁴⁸

2.78 Another opinion from a former academic takes this up:

What is happening is a diminution of standards, a negation of the concept of excellence—this one-size-fits-all model that says that nobody will fail, we'll all be happy, and we wouldn't want to hurt anybody's self-esteem by saying that they could work harder and improve.⁴⁹

2.79 The committee supports plain English report cards as the best way to inform children and parents of academic achievement and progression.

Parental concerns about reporting

2.80 The committee received some submissions from parents who were highly disappointed with their child's levels of achievement. This disappointment was heightened by the relevant school's failure to adequately inform the parent of how his or her child was progressing.

2.81 The Year 7 or Year 8 teacher will have the task of dealing with low-performing students while catering for high-achieving. An inexperienced teacher can fail at both ends of the scale.⁵⁰ One parent submitted that she had been misled by a reporting practice which was verging on dishonesty:

My son has attended our local Catholic primary school since Prep. The school kept sending home good reports and awards that told me my son was progressing and these reports have been disguising the fact that my son has not learnt to read. My son is 12 years old and has a reading age of just 6.2 years, according to several educational psychological assessments. He is therefore 6 years behind, still at a Grade Prep/1 level when he actually is in Grade 6...My son faces high school in 8 months at a very shocking pre-school standard.⁵¹

2.82 Another parent, Yvonne Meyer, provided the committee with another instance of how parents may be misinformed:

People think words mean one thing, and they do not; they mean something completely different—such as being fobbed off with these overly optimistic

48 Mrs Valerie Gould, Association of Independent Schools of Western Australia, *Committee Hansard*, Perth, 2 July 2007, p. 3.

49 Professor Stephen Kessell, *Committee Hansard*, Perth, 2 July 2007, p. 62.

50 Ms Joy Schultz, *Submission 59*, p. 1.

51 Private Submission, *Submission 23*, pp 1-2. Also, Private Submission, *Submission 29*, pp 1-2.

school reports. Few parents realise, for example, that here in Victoria, in year 12, the kids are graded across nine levels, from A+ all the way down to E, essentially, although they do not call it that. C is in the middle. C should be the average grade. Yet the most commonly awarded grade at year 12 is A. So in fact A is average, A+ is above average and B is average. So, if a child comes home with a B, the parent thinks, 'Well, that's pretty good,' because one assumes that C is average and a B is above average. It is only when parents are told that 35 per cent of students in year 12 are awarded an A that suddenly the meaning becomes apparent. But parents are not told this.⁵²

2.83 The point of this is that information to parents on the progress and achievement of their children should be readily comprehensible and adequately convey whether a child is progressing as well as might reasonably be expected. The committee could not say precisely what form of reporting would best serve the needs of parents and students except that there was general agreement that current reporting terminology is inadequate. There is often confusion about whether marks and grades are given on the basis of criterion referencing or normative referencing. The distinction should be made clear to parents, and other interpretation explanations given on the reports. This is a responsibility for school systems, and possibly state boards of studies as well. The following comments confirm the committee's concerns:

The provision of a ranking on some graded or numerical scale [fails to] give parents the kind of information they really want...It also has the potential to lead to unrealistic expectations...The essence of feedback to parents must be descriptive.⁵³

2.84 If school principals believe this issue remains a problem after so many decades of reporting, it is time that some serious research-based policy be determined. The committee also understands the importance of reporting on the overall growth of a student, as expressed below:

The current accountability requirements are perceived to be onerous and make significant additional demands on teachers' time. Assessment should be beneficial to students' learning and the reporting of achievement should be informative to their parents. The norm-based standards of assessment...only focus on a very limited aspect of the assessment of learning. Students need to be given the opportunity to demonstrate their knowledge and understanding in a variety of ways.⁵⁴

52 Ms Yvonne Meyer, *Committee Hansard*, Melbourne, 25 June 2007, p. 56.

53 Australian Primary Principals' Association, *Submission 43*, p. 14. Also, Dr Glenn Finger et al, *Submission 46*, p. 6.

54 Lutheran Education Australia, *Submission 41*, p. 6. Also, Mr Chris Watt, Independent Education Union of Australia, *Committee Hansard*, Melbourne, 26 June 2007, p. 13.

2.85 Another submitter strongly criticised the Queensland assessment systems for being vague, wordy, undefined and dependent on an 'overall judgement'.⁵⁵

2.86 However, the most confusing method of reporting students' results was described at the committee's hearing in Perth. In Western Australia, the committee was told:

The government sector has now set targets for years 3, 5, 7 and 9 so that, if students get a level 2 in year 3, they will be given a B; if they get a level 3 in year 5, they will be given a B; and so on as it goes up. Because the levels are quite broad, it actually divides those levels into three bands—first, middle and high. It may be that you are part of the way through level 4 in year 7 to get a B but you have to be all the way to the end of year 4 and year 9 to get a B. They have been aligned against the levels and the levels are clearly defined. Teachers will make judgements on what level the student is at and then, depending on the year of schooling, an algorithm will tell you if you are an A, B, C, D or E student.

Basically saying that if you have all level 3s and above in year 5 you would be a B student, but if you had some level 4s in year 5 you would probably be an A student. It is about how many level 3s or 4s you have according to the year. If you got a level 4 in year 3, you would be an A student. If you got a level 4 in year 5, you would be an A student. If you got a level 4 in year 7, you would be an A student. But if you got a level 4 in year 9, you would only be a B student.⁵⁶

2.87 With due deference to the experienced teacher who is the witness quoted, the committee has only a hazy understanding of what this all means, even after several readings of the Hansard. That itself is a matter of concern. As described in a later chapter of this report, Western Australia is recovering from a prolonged bout of outcomes based education, and this may be part of a residue of policy which remains to be swept away. It serves, however, to illustrate the tension between the need to report progress to parents in an intelligible way, and at the same time to ensure that assessment of achievement is carried out in a way which accords with the best teaching and learning practice. The committee understands that there will be problems in negotiating something that gives due weight to concerns on both sides.

2.88 The Commonwealth has insisted that states and territories report to parents about student progress on an A-E scale. This has caused problems for Western Australia, as explained above. One example of the problems caused by the Commonwealth requirements was described by Professor Loudon, now head of the Curriculum Council in that state.

Local teachers are struggling trying to find a way to match the federal government's desire to have every children [sic] get an A, B, C or D, which

55 Dr John Ridd, *Submission 4*, p. 11.

56 Mrs Valerie Gould, Association of Independent Schools of Western Australia, *Committee Hansard*, Perth, 2 July 2007, pp 8-9.

is a funding contingent issue for the state government. The state government does not believe in it...So they have very highly elaborate ways of generating marks which then get converted. My view, as it happens, is that the federal minister was right to pick out talking to parents that they found that our Australian reporting system is obtuse. They could not figure out what they meant and they were full of words and words. The community view was to just give them a mark.⁵⁷

2.89 This was then complicated by the awarding of an A grade to the students who achieved the benchmark level set:

I would have thought that an A grade would have been better delivered to students who are a number of bands above the minimal standard. That is where I think the system here fell apart with the grades.⁵⁸

2.90 While the reporting might be against the standards, not every parent in Western Australia will be informed about levels and bands. Perhaps this is why the independent schools in Western Australia have in some instances reverted to percentages. Not only does this peculiar reporting method significantly increase teachers' administrative workloads, it might also be counter-productive for those children who are the lower performers or disengaged with education.⁵⁹

2.91 The committee emphasises that while the problems in Western Australia are not found elsewhere, they illustrate a point of tension in reporting that is felt much more widely. It is also hoped that these tensions in the west will fade as policy is revised.

Conclusion

2.92 The committee might be reassured by the results of the PISA and TIMSS tests, which put Australia toward the top of all but the highest category of performance, but it believes that there is a warning in the existence of a long tail of underperformance. It notes also that Canada, a country with many points of commonality with Australia, has the same performance but without the tail. In the next two chapters of the report, education quality issues will be discussed in such a way as to explain why this tail exists, and what can be done to shorten it.

2.93 On the more immediate issues discussed in this chapter, the committee is concerned that benchmark testing, which it supports, is not being taken up more enthusiastically by schools. It notes the reasons why this is so, and **recommends** that

57 Professor Bill Loudon, *Submission 73*, p. 7.

58 Professor Michael O'Neill, University of Notre Dame Australia, *Committee Hansard*, Perth, 2 July 2007, p. 44.

59 Mrs Valerie Gould, Association of Independent Schools of Western Australia, *Committee Hansard*, Perth, 2 July 2007, p. 10.

efforts be made to give the tests more credibility and usefulness as teaching instruments.

2.94 Finally, the committee notes the continuing argument over reporting. While it believes that the A-E scale carries much more meaning for parents than other systems that have been in use, it is time to examine more closely the need for information to be provided which explains students' results and where students are achieving relative to others. The use of performance indicators should give parents an honest view of how their children are performing against the standards.

Recommendation 1

The committee recommends that efforts be made to give the national benchmark tests more credibility and usefulness as teaching instruments.

