Assessment and Reporting for the middle years fall into the structural ‘middle’ gap between the early years of beginning primary school, and the final years of completing secondary school. They also fall into the gap between traditional assessment methods, such as formal examinations and projects, and evolving newer methods of diagnostic assessment. To understand these two kinds of gaps it may help to consider some background.

Assessment is the major focus in two of the five fundamental, interrelated, equally important components of the education process:

- Finding out what students already know about some topic
- Choosing, making, preparing what is to be taught
- Teaching it
- Helping students learn what has been or is being taught
- Assessing how well the students are learning this.

But amongst these, assessment has had an uneven history.

In the mid-1970s a national scandal erupted with the publication of literacy and numeracy test results on 10-year-old and 14-year-old students. Approximately one-quarter of the 14 year-olds tested were unable to read the time on a clock-face with hands, in an era when digital clocks were a rarity, and analogue clock-reading was an essential life-skill. The resulting backlash against testing, nationally and across
Australia’s states, had a lasting impact on school assessment, and teacher-training. Formal examination systems that had dominated assessment at all levels of secondary schooling, with parallel equivalents in end-of-term and end-of-year tests in upper primary grades, were swept aside.

Despite this, in the last one or two years of secondary school it was still necessary to select as effectively as possible the best school-leavers for entry to scarce, expensive tertiary courses, vocational training, and employment. Formal examinations remained the standard method although, increasingly, non-exams such as projects contributed to students’ final subject scores and tertiary-selection ranking. This summative end-of-secondary assessment is still a major feature of assessment in schools. Necessarily students should prepare earlier for this, to develop exam-taking skills, and avoid excessive exam-anxiety.

At lower year levels, examinations are currently less important. But concern for standards remains—how well are students learning literacy, numeracy, and other curriculum areas? One major global stimulus to such concern is the recurring international assessment and comparison of schools in mathematics and science, such as Third International Mathematics and Science Study (TIMSS). Another stimulus has been the successive Hobart Declarations (1989, 1997) and the Adelaide Declaration by the Australian Education Council (AEC), now Ministerial Council on Employment, Education, Training and Youth Affairs (MCEETYA), which led to draft national profiles of curriculum, along with state-based alternatives.

Recent developments have led to a revaluing of assessment, along with the introduction (and, in part, re-introduction) of state-wide testing of literacy and numeracy at Years 3, 5, 7 and 9. Little that is being proposed and/or implemented now is genuinely new, although the vocabulary and supporting theories are superficially different from past approaches. However, new or not, the decades old assessment vacuum is being refilled. This article outlines some of the seemingly new developments.

Rubrics

One topic seems new; ‘rubrics’. More than 10 years ago they did not exist, except in old prayer-books, as a margin-note advising when to stand, kneel, or do some other action. Now rubrics are all the rage (NCTM 2003, Stenmark et al. 2001). What are they? Typically, in this new sense a ‘rubric’ is also a table, with rows and columns. The columns identify the level or quantity of achievement, understanding, or skill—ranging across:

- Not Begun and beginning
- Developing or consolidating
- Established
- None or not very much, or rarely
- A little, or seldom
- Quite a lot or often
- Usually or large amount.

The rows identify subject-related aspects of what was being learned. That is, a rubric is essentially an elaborated, tabulated checklist of expected learning outcomes, with
graded exemplars of observable behaviours. It enables a rubric-user to identify that Student X has learned Objective P to Definable/Observable level D. A rubric is a weighted, or quantified checklist. Although this use of ‘rubric’ is new, what it means is not new for actual teaching and assessment practice. Consider the graded criteria for assessing Year 12 projects that teachers use: these are essentially rubrics for grading project performance (Gough 2006:8-9).

During the heady 1970s, not only was formal assessment largely absent, but curriculum was often school-based, after decades of centralised prescription by official syllabuses or textbooks. Inevitably it became necessary to reconsider such a free-market approach. The result was that, around Australia (and overseas) existing curricula were surveyed, summarised, and redefined in terms of more loosely indicative (not prescriptive) frameworks of outcomes or objectives. This was essentially returning to the behavioural objectives that had dominated the 1960s, following the lead of Benjamin Bloom and colleagues (Bloom et al. 1956). The latest incarnation of this (Anderson & Krathwohl 2001), renews emphasis on higher level thinking skills of analysis, synthesis, problem solving, problem posing, creativity, and communication.

The New Basics of Queensland, for example, with its so called Productive Pedagogies, and the Victorian Essential Learning Standards (VELS 2005) and the earlier Curriculum and Standards Frameworks (CSF II: Board of Studies 2000), in their different ways, attempt the same thing: to specify what teachers might teach. Given some detailed specification, it follows logically that teachers will check how well a student has learned from the classroom experience that was developed to teach that specification. A rubric is an organised collection of specified curriculum outcomes and graded performance descriptors—little more than a useful common-sense tool for identifying and describing how students are progressing.

FORMATIVE ASSESSMENT
More interestingly, while summative assessment has retained its usefulness in the last stage of secondary schooling, formative assessment has been researched, and shown to be highly effective as a way of promoting learning at all levels of schooling: it is arguably assessment for learning. (Summative assessment, investigates, analyses, describes, and judges what has been learned, without seeking to offer constructive advice about what to do next, or what to do if learning is found to be weak or patchy. Formative assessment offers diagnostic advice, shaping or forming the follow-on curriculum.) Research by Paul Black, and colleagues of the Cambridge University (School of Education) Assessment Reform Group shows that if teachers assess formatively, and use the assessment results to shape their subsequent teaching, negotiating with students what the students should do to learn more, then, not surprisingly, the students do learn.

This is, of course, dramatically different in its use of assessment (especially testing), compared with traditional approaches that accumulated week-by-week spelling and arithmetic test-scores, along with project marks or topic tests, and term tests, and added everything together to establish that, overall, Ferdie had a final score of 83 per cent in English, and a final score of 57 per cent in mathematics, whereas Mortie had scores of 74 per cent and 68 per cent respectively. By receiving diagnostic formative
advice about that relatively low-scoring end-of-April writing assessment, Ferdie might be helped to develop greater skill with written expression and proof-reading, while similar advice about weak test results on multiplying and adding fractions might help Morrie improve in that area.

An important and genuinely new development in assessment has occurred in Victoria, with its Early Years Numeracy Interview (for example Clarke 2000; Early Years Branch 2001), and in and other states, such as Queensland, with its Diagnostic Net (Paxton, Wolfe, & Zevenbergen 1998). Based on restated curriculum frameworks for primary mathematics, in the case of the Early Years Numeracy Interview, a one-to-one interview is conducted with school-beginners, about five years old, establishing how much of the proposed mathematics curriculum they already know. Not only are they asked questions that would normally be part of the curriculum for that first year, other, later, harder questions are asked, as long as the child is able to answer the questions correctly. Only when a child can no longer answer questions on one topic is the topic changed (for example, from counting with whole-numbers, to questions about reading and using a clock) and the interviewing resumes, with early, easy questions, progressively working through harder and later questions.

Although this is not a new idea, making a start-of-year diagnostic interview statewide policy is new. It is radically different from the usual approach to assessment, using questions that are strictly at the presumed level of difficulty for the students being assessed. Typically, for students in Year 3, or entering Year 3, for example, the questions would come from the more or less standard Year 3 curriculum (Schleiger & Gough 1993). By contrast, the new ‘diagnostic profile’ approach deliberately includes questions that start early and easy—around Kindergarten or Preparatory—and become progressively harder, up to about Year 8 level.

At the end of an Early Years Numeracy Interview, or an equivalent diagnostic profile (such as the follow-on Middle Years Numeracy Interview developed by Siemon and colleagues in 2000), the teacher knows what the student already knows and can do. The teacher then prepares suitable curriculum materials and learning experiences that will help that student learn more, starting from where the student is.

In the area of literacy, such an approach would be equivalent to starting the learn-to-read (or write) curriculum by having the teacher work individually with five year olds, finding out who can already read (and/or write), and how well they can do so. Then, of course, the teacher proceeds to work with the students individually, or in small groups, teaching further reading (or writing) skills, and language study.

RICH ASSESSMENT TASKS
And middle years? What is missing, as far as I am aware, is any counterpart to the Early Years Numeracy Interview, apart from my own Diagnostic Mathematical Profiles (Gough 1999, based directly on Schleiger’s earlier Diagnostic Mathematical Tasks, which were year-leveled). Outside mathematics, little seems to be available to use in such a ‘diagnostic profile’ way.

However, a very different approach to assessment could be developed to help fill this large gap. Given the obvious need to establish, through the middle years, what individual students already know about a particular curriculum area, to start their
Further learning at that point, open-ended questions are useful. These are loosely related to the idea of ‘rich assessment tasks’ or ‘RATs’. Interestingly, decades earlier John Biggs proposed the use of potentially rich learning situations (with the acronym PRS), an early anticipation of the move to invigorate the teaching of problem solving in all curriculum areas (Biggs 1975). Importantly with PRSs and RATs individual students begin working at their own level of knowledge or skill, and pursue the task or question in widely differing ways.

Open-ended assessment is comparatively easy with literacy. A teacher can prepare a set of books (novels, non-fiction, plays or poetry; or pages from such a selection), ranging in difficulty from very easy to rather hard (for example, from simple picture-story books to Robinson Crusoe or War of the Worlds, in the originals). Then the student can be shown, one at a time, a page from each, and asked to read the page aloud. Judging whether or not the child can do this adequately is easy enough, and indicates the level of reading-difficulty at which the student would usefully re-commence reading instruction, language study, and literacy practice. Similarly, simply by asking a student to write a story, or a letter, it is possible to gauge the student’s level of skill in written expression. Open-ended assessment in other curriculum areas could be based on questions which search for how much the student knows or can do: asking a piano student to play the hardest piece he or she knows, or asking a science student to describe the most difficult piece of science he or she knows.

Diagnostic profiling is one of the main ideas in Roy Killen’s recent book on programming (curriculum choosing and construction) and assessment (Killen 2005). Killen outlines effective assessment methods that incorporate the current ideas of ‘Quality Teaching’ (NSW), ‘Productive Pedagogies’ (Queensland), general high-level outcomes (promoted by Spady 1994, and Mayer competencies 1993), and other learning outcome approaches (Victoria and Tasmania’s essential learnings), as well as the national benchmark system (Australian Benchmarks 2005).

**PORTFOLIOS AND AUTHENTIC ASSESSMENT**

Portfolio assessment is a newcomer (Forster and Masters 1996, Watson 2002). Portfolios are ideal as a gallery for displaying either developing drafts of learning, or selections of best achievement. However as teachers’ own use of professional portfolios (as an extended counterpart to resumes or curriculum vitae) is well known, along with the serious difficulties in making a consistent, objective summary of such portfolios, no more needs to be said here, except that the same problems apply to the potential diversity of student portfolios.

Only one further type of assessment, that might be new, or might be relevant, need concern us here—so called ‘authentic assessment’. The very name is problematic, because its honorific name threatens to condemn any other kinds of assessment as ‘in-authentic’. The technical definition of ‘authentic assessment’ is that it is based on so called ‘authentic activities’, that aim to achieve real-world, purposeful, practical goals (Anderson, Rede, & Simon 1995). What must be stressed, here, is that, valuable as many purposeful (non-recreational) real-life tasks are, as topics for classroom (and out-of-class) experiences, many otherwise ordinary curriculum topics cannot be justified as ‘authentic’ in this sense, but deserve to be presented...
and assessed as seriously as ever. Consider, for example, quadratic equations (or any other medium for developing robust algebraic skill), history, oil painting, chess, and sonnet writing—things that arguably are, for all practical purposes, pointless, yet are still fundamental aspects of human culture. Hence, ‘authentic assessment’ should not be misunderstood, and should be treated cautiously. It is not a panacea or a total replacement.

ACKNOWLEDGEMENTS
Thanks to Judy Mousley for advice about an early draft of this discussion, and to the stimulus of teaching assessment issues to many students.

REFERENCES


