

Standards as central to improvement

Currently there are clear signs of a concerted move towards National Curriculum, bringing with it much debate about domain knowledge and the related capabilities to be developed in students. This is a time of historic significance in the sense that the moves to National Curriculum are occurring in the first year of National Testing for literacy and numeracy in years 3, 5, 7 and 9. All of this is occurring against a backdrop of assessment writing showing an increasing interest in assessment in a learning culture (Shepard, 2000), with quality teacher assessment understood as central to local and system efforts to improve student learning and outcomes. Also of note is international concern about the utility of testing for monitoring standards of achievement in schools, which has led to renewed interest in the links between standards, school-based teacher judgements and quality learning.

Broadly speaking, such interest has concentrated on formative assessment for diagnosis and improving learning, with summative assessment for reporting purposes remaining the province for more centralised control at system level. These related matters raise a suite of issues around the nature of quality assessment, the factors that underpin and motivate how assessment is developed and enacted, how the option of teacher assessment for summative purposes can be adopted with confidence, and how we understand, interpret

and use the evidentiary base that assessment practices call forth, in system and local school contexts.

While these issues are priorities in themselves, their importance is further raised by the public policy in education striving to align assessment and reporting in ways that will have profound effects on learning and teaching. This article is based on three main propositions:

1. that there is considerable international interest in undertaking rigorous examination of how teacher-led, quality-assured assessment judgements could be used to generate assessment data for system-wide accountability and reporting purposes
2. that curriculum reform in educational policy, in and of itself, will not lead to improved student learning
3. that efforts to realise *improvement for all* need to focus on *quality* and defined achievement *standards*.

While an A to E reporting framework has been initiated across the country in an attempt to provide clear and consistent reporting, clearly defined statements of expectations (i.e. standards) at each of these levels have not been developed to date. This means that the potential of the framework to improve reliability of teacher judgement has not been realised. Further, the reporting framework falls short of supporting efforts by education systems across the country seeking to use classroom assessment evidence for system reporting and for tracking achievement over time.

Currently in Australia matters of curriculum, assessment and reporting are under intense scrutiny, attracting high policy interest and considerable investment. CLAIRE WYATT-SMITH contributes to current debates about the nature and purposes of assessment.

The key link that needs to be explored in Australia is assessment standards and quality: specifically, the centrality of standards to system efforts to provide clear communication about what students are learning and achieving at school. While both international testing programs (e.g. PISA) and census standardised literacy and numeracy testing programs have been installed, currently there is little evidential base about student achievement in English, science and mathematics in which education systems, teachers, parents and the wider community can have confidence. There is a well-recognised need to provide such evidence, contributing to the growing call for a national and state assessment culture with standards at the centre. The potential of achievement standards lies in their location at the interface of assessment practices at the local level and the central political demand for public accountability at system level. A key driver for this move is that quality standards can be used to provide much-needed evidence to inform decisions at both levels—local and system.

To have confidence in decisions at both levels, it is important to achieve high reliability while preserving validity in the use of standards for assessment. Several writers have argued that to achieve such confidence it is important for teacher assessors to develop common understandings of mandated standards and reach 'similar recognition of performances

It is this final call for student involvement in assessment practices that has the greatest potential for improving student achievement and requires teacher knowledge of standards and how to use them for improvement purposes.

that demonstrate those standards' (Maxwell, 2001, p. 6). However, clear communication about the nature of standards and the levels they seek to specify is not necessarily achieved through the provision of stated standards alone. Sadler (1989) argued that exemplars or samples of student work provide concrete referents for illustrating standards that otherwise remain abstract mental constructs. He made the point that the stated standards and exemplars work together to show different ways of satisfying the requirements of, say, an A or C standard. A related point is that standards written as verbal descriptors call for qualitative judgements. As such, the standards remain open to interpretation and common understandings of the terms used to capture notions of quality in the standards need to develop over time.

My work in assessment in English curriculum in secondary schooling showed that exemplars can include annotations highlighting features of the work showing how these match the stated criteria (Smith, 1989). More than this, however, carefully chosen exemplars illustrating the standard at mid-band and threshold levels need to include an accompanying commentary showing how the teacher assessor matched the student work with the stated standard, and the cognitive processes this entailed in combining strengths and limitations of the work related to the standards. In short, annotated exemplars, each with an accompanying commentary, can make explicit the judgement processes involved in arriving at an overall grading decision. In the absence of such a commentary, the

treatment of compensatory factors and the complex features of teacher judgement remain unarticulated, leaving no trace in a grade recorded on a student piece of work.

While standards and commentaries can serve to make clear expectations of quality, they do not necessarily account for the factors that shape teacher judgement. In a large scale Australian study of teacher judgement in middle schooling, Cooksey, Freebody and Wyatt-Smith (2007) reported high levels of variability in teachers' notions of quality and also unearthed a range of factors that shape how judgements are reached. While this study pointed to the need for the promulgation of stated standards to include exemplars, it also opened a vital space for considering social moderation as focal in quality assurance processes at local and system levels. Specifically, it suggests how social moderation can act as a context for teachers to make available for scrutiny the bases of their judgement practices and their use of standards in those practices.

However, as Wyatt-Smith and Gunn explain, 'the reality is that while many teachers have initiated their own professional conversations around assessment practice, both within their school and at district level, it is also fair to say that many teachers experience a sense of isolation as they go about their work as assessors, having no sustained opportunities for such sharing'. They argue for a 'pressing need to support teacher dialogue (i.e. through social moderation) around the issues of assessment and judgement, including standard setting and how to make available for students useful information about expectations of

quality'. It is this final call for student involvement in assessment practices that has the greatest potential for improving student achievement and requires teacher knowledge of standards and how to use them for improvement. In conclusion, an assessment system based on quality and defined achievement standards that places the masterful teacher at the centre has much to offer at local and system levels. ☺

References

- Cooksey, RW, Freebody, P, & Wyatt-Smith, CM (2007). 'Assessment as judgment-in-context: Analysing how teachers evaluate students' writing' in *Educational Research and Evaluation*, vol. 13, no. 5, pp. 401-34.
- Maxwell, G (2001). *Moderation of assessments in vocational education and training*, Department of Employment and Training, Queensland.
- Sadler, DR (1989). 'Formative assessment and the design of instructional systems', *Instructional Science*, 18, pp. 119-44.
- Shepard, LA (2000). *The role of assessment in a learning culture*, Presidential address, American Educational Research Association, New Orleans, April 2000.
- Smith, CM (1989). 'A study of standards specifications in English', unpublished masters thesis, University of Queensland, Brisbane.
- Wyatt-Smith, CM & Gunn, SJ (in press). 'Towards theorising assessment as critical inquiry' in CM Wyatt-Smith & JJ Cumming (eds), *Educational Assessment in the Twenty First Century*, Springer, Dordrecht.



Claire Wyatt-Smith is dean of the Faculty of Education at Griffith University in Brisbane.

