

**Quality management of school-based assessments:  
Moderation of teacher judgments**

Graham Maxwell

Paper presented at the 32<sup>nd</sup> IAEA Conference, Singapore, May 2006

Abstract

*School-based assessment offers many benefits over external tests and examinations. These benefits include attention to a greater range of important learning outcomes, greater integration of formative feedback for improvement, and generating a performance profile over time rather than on a single occasion. These fit with current understandings of human learning and with anticipated future demands on people and economies.*

*A critical issue is how to establish confidence in school-based assessments. Greater confidence, and therefore stronger quality management, is needed for higher stakes assessment.*

*A key component of successful school-based assessment is teacher expertise. Two kinds of expertise are involved: obtaining good information on student learning (using good assessment procedures); and making good judgments (applying relevant performance standards). Inservice teacher education programs are essential and exemplars of good practice can be useful. Self-monitoring (quality assurance) processes can be useful. However, for high-stakes assessment, some form of external quality control (moderation processes) is essential.*

*This paper discusses various options for moderation processes and their advantages and disadvantages. Reference is made to Queensland's three decades of experience with school-based assessment in secondary schools.*

In this era of rapid change, new understandings are emerging about the nature and process of human learning. Some of this comes from research on learning processes and some from neurological research on the brain (e.g., see Bransford, Brown & Cocking 2000). These new understandings affect the way we think about the quality of learning, its development over time, the scope of that learning, and the kinds of outcomes that we should expect. They also affect the way we think about assessing that learning, away from the narrow confines of standardised tests with their discrete and decontextualised 'items' towards more complex, wholistic, contextualised and authentic forms of assessment (e.g., see Pelligrino, Chudowski & Glaser, 2001).

As well, the demands of the global economy are becoming more insistent, especially in terms of the increasing value of intellectual capital, creative imagination and the application of knowledge. Citizens of the future need to be flexible and practical problem solvers and capable of life-long learning.

In this context, school-based assessment offers many benefits over external tests and examinations. These benefits include attention to a greater range of important learning outcomes, opportunity for contextualised and authentic assessment, integration of formative feedback for improvement, and generation of an achievement profile over time rather than on a single occasion. These fit with current understandings of human learning and with anticipated future demands on people and economies.

The inclusion of a greater range of learning outcomes in assessment is an important issue. School-based assessment can include practical and performance assessments (e.g., projects, designs and presentations) that cannot be included in external assessments because they are too time-consuming or context-dependent. There can be a more deliberate fit between learning expectations, learning support (teaching) and assessment.

This does not mean that there needs to be a contest between school-based and external assessment. There can be complementary roles for each. Each can deliver different benefits — especially, greater validity from school-based assessments and greater reliability from external assessments — though these benefits are possibilities rather than certainties. Bad practice negates any benefits.

There is also an issue of which takes precedence where both are practiced. Typically, external tests take precedence because they are more ‘trusted’. The question is how to grow more trust in school-based assessments. Essentially, if school-based assessment is to play a less subservient role, then the quality of those assessments needs to be managed. But how much quality management is needed? That depends on whether the assessments involve high or low stakes. The higher the stakes, the greater the need for confidence in the outcomes and therefore the stronger quality management needs to be.

A key component of successful school-based assessment is teacher expertise. Two kinds of expertise are involved: obtaining good information on student learning (using good assessment procedures); and making good judgments (applying relevant performance standards).

### **Quality management**

A quality management system can have different components covering the periods before, during and after the assessments take place. Although there is no common terminology, one way of distinguishing different components is to say that quality assurance occurs before, quality control occurs during, and quality review occurs after the assessments are completed and the results reported. In this sense, quality assurance is concerned with establishing appropriate circumstances for assessment to take place, quality control is concerned with verifying that assessment procedures and judgments are appropriate (and

'approving' them), and quality review is concerned with retrospective analysis to see whether improvements in assessment procedures and assessor judgments should be made in the future. In this paper the focus is on quality control.

One approach to quality control is referred to typically as moderation. This, here, refers to a process of quality control involving the monitoring and approval of assessment procedures and judgments to ensure there is consistency in the interpretation and application of the performance standards. This can involve a single person (a moderator) or a group of people (a moderation panel) looking at samples of evidence of student performance and determining whether they agree with the assessment judgment of the assessor. Feedback to the assessor may include advice on improving their assessment procedures or evidence gathering and adjusting their assessment standards, that is, modifying their judgment of the standard of performance. Moderation can have a collegial orientation of support rather than a managerial orientation of control. Every moderation system needs to determine where the ultimate authority lies for approval, what is an appropriate balance of rights and powers of all participants, and how differences of opinion can be resolved.

'Moderation' usually means 'social moderation' in the sense used by Linn (1996). That is, it involves comparisons of the assessment judgments of different assessors in different settings but all relating to the same learning outcomes (or competency standards), with the purpose of ensuring that the judgments are comparable. Comparability of assessment judgments means that there is agreement that the assessed performances are appropriately classified in terms of the standard they demonstrate. This involves both similar interpretation of the standards and similar recognition of performances that demonstrate those standards. A moderation process is therefore one involving approval of assessor judgments, with the implication that there may need to be some adjustment of those judgments to conform to the common standard. It is not a passive process that simply checks how much agreement there is but an active process in which assessment judgments are aligned with each other to create consistency of interpretation and implementation of standards across the whole system. Differences of opinion therefore must be resolved rather than simply noted and accepted.

Moderation systems can differ in style and complexity. Four basic approaches can be identified, under two general headings, as shown in table 1.

**Table 1: Types of moderation systems**

<b>Strong control; high stakes</b>	<b>Weak control; low stakes</b>
External moderator	Assessor meetings
External moderation panels(s)	Assessor partnerships

Strong control is necessary where the assessments involve high stakes. Weak control is more appropriate where the assessments are low stakes. 'High stakes' refers to situations where the consequences of the assessment can be considered serious (for the student or the school). 'Low stakes' refers to situations where the consequences are less serious. Clearly, the distinction is not categorical; rather, there is a gradation from higher to lower stakes and the distinction is somewhat subjective. Decisions about whether the stakes should be considered high or low requires judgment of the seriousness of the consequences in particular circumstances.

### **External moderators**

External moderators offer strong external control. 'External' here means external to the school. The external authority resides in a certifying agency that is therefore responsible for confirming or approving the assessments. The moderator is trained for the role and will assume responsibility for moderation of an appropriate number of assessors. This can be defined as a geographical area. There could be a number of external moderators for the state. Moderators' activities typically involve review of student portfolios as well as site visitation for discussion with assessors and observation of assessment situations. Clearly they need to be experienced in training and assessment in the relevant qualifications.

The advantages of external moderators include:

- they can offer authoritative interpretations of competency standards
- they can carry the standards from site to site and assessor to assessor
- they can offer advice on assessment approaches and procedures
- they can observe actual assessments not just view folios
- they can be a trouble-shooting resource for assessors to draw on
- they can induct novice assessors quickly into high quality assessment.

The disadvantages of external moderators include:

- there are substantial costs involved (salary, base office, travel accommodation, communications, training, moderator conferences, etc.)
- there are logistic problems to be overcome in covering all assessors adequately
- 'authoritative' interpretations are not always right or appropriate
- 'external authority' can be stultifying rather than liberating, encouraging conformity rather than innovation.

### **External moderation panels**

External moderation panels offer another form of external confirmation of assessment decisions by a certification agency. This is the approach taken in Queensland senior secondary schools where the certification agency is the Queensland Studies Authority (QSA) and the student assessment is wholly

school-based. Although external moderation panels can function in a variety of ways, what will be described here is the way that QSA panels currently operate.

QSA moderation panels, called review panels, consist of selected teachers, chosen for their experience and excellence in teaching and assessment. Their participation is voluntary and supported by their school. Most would see participation as a valuable professional development opportunity. It does not involve additional remuneration and the cost of involvement is borne by the schools. Panel membership is reconstituted every six years, involving replacement of at least one-third of the panel.

The state is divided into administrative districts. In each subject or subject area, there is a district review panel in each district plus a state review panel; some subjects with small enrolments have different arrangements. State review panels moderate the operation of the district review panels, advise on any unresolved issues within the district review panels, and resolve disagreements between schools and district review panels.

The moderation process is seen as consisting of the following components:

- accreditation of school plans for implementing the subject syllabus (a form of quality assurance)
- review of each school's assessments through monitoring, verification and approval (see later)
- random sampling (a post-hoc mechanism for evaluating the effectiveness of the moderation process).

The review process forms the heart of these moderation procedures. Monitoring advises schools on the appropriateness of their assessment processes and standards at the end of year 11 (half way through the course). Verification advises schools towards the end of year 12 on the appropriateness of their judgments of standards of performance of their students. Approval is the final process of validation of the results to appear on the Senior Certificate.

The part of this process that is of most relevance here is the operation of the review panels during monitoring and verification. A sample of student folios is sent from each school in each subject to the relevant review panel. Each folio includes the relevant pieces of student work or other records of their performances as well as the school's judgments of the standard reached by the student. Teachers' judgments concerning how close each student is to the grade boundaries are also provided. Adjustment for any student in the sample can have repercussions for other students in the group.

Members of the review panel meet and review each of the sampled folios, considering whether they agree with the teacher judgments of standards demonstrated by each student. The panel seeks agreement within the panel before offering its advice to the school. Where adjustments are recommended,

the panel chair may enter into discussion and negotiation with the school. A process exists for appeals. Ultimately, the QSA has the power to reject a school's opinion. However, disagreements are negotiated at length and in good faith and almost always satisfactorily resolved.

There are, of course, many other features to the QSA moderation process but these are the details most relevant to the issues being addressed here. Further details are found on the QSA website ([www.qsa.qld.edu.au](http://www.qsa.qld.edu.au)).

It is important to notice the following features of this panel-based moderation:

- the initial process involves a quality assurance approach (accreditation)
- the assessment judgment is a global judgment, that is, a synthesis of all the performance evidence in the student's portfolio
- advice is offered to schools mid-course not just at the end, which allows schools time to make adjustments to their assessment procedures and standards
- a sample of student portfolios from each school is considered, not all students (except where there are very few students)
- review panels consider whether they agree with the school's judgments of student achievement in full knowledge of those judgments (that is, the focus is on confirming or challenging the school's assessment judgments, not on making an independent or 'blind' judgment which is then compared with the school's)
- advice is offered to schools on the quality of their assessments and the appropriateness of their judgments of standards
- disagreements between a school and a panel are discussed and negotiated
- QSA has ultimate power of approval and some things may be non-negotiable.

Advantages and disadvantages of external moderation panels depend on how they are implemented. However, there are some general characteristics, especially when contrasted with external moderators.

The advantages of external moderation panels include:

- panels are likely to make more consistent decisions than individuals
- panels can cover a larger group of assessors than a single moderator can
- panels can offer more comprehensive advice on assessment approaches and procedures (by being able to draw on a wider cross-section of examples)
- panels represent collective authority rather than single person authority
- they provide powerful professional development for those involved.

The disadvantages of external moderator panels include:

- there are substantial costs involved (travel, accommodation, communications, training, moderator conferences, etc.)
- there would need to be some fulltime officers to organise the process (so

- the costs would not be less and may be more than for external moderators)
- there are logistic problems to be overcome in covering all assessors adequately
- panel activities take members away from their other work (and may require payment).

### **Assessor meetings**

Assessor meetings require the participation of all assessors. These may occur within an institution or across several institutions. In Queensland senior secondary schools, it is assumed that there will be meetings of all teachers of a subject within a school at various stages of teaching and assessing to establish a process of within-school moderation. These stages might be before beginning teaching (to align their expectations of students), before undertaking an assessment (to discuss the criteria and standards for assessing performance), after undertaking an assessment (to discuss whether they have applied similar standards and to make adjustments where necessary) and at the end of the course (to compare and confirm their final judgments of the exit standards reached by students).

At times also there have been meetings of teachers from several primary schools to compare their assessments of students and moderate their judgments against common benchmarks of learning outcomes. Clearly, under these conditions, the incentive to participate and to modify one's own practice depends on professional motivation not on any sanctions. Even if such meetings are mandated, unless there is a certification function to which they contribute, personal engagement is dependent on teachers seeing that the meetings are helpful for improving their own practice rather than simply verifying whether they are adopting common standards. Again, though, this can be a powerful agency for raising the quality of teacher practice.

The advantages of assessor meetings include:

- the opportunity for direct comparison and sharing among assessors
- a less judgmental atmosphere than for external moderators or panels
- personal ownership of any new insights and understandings and ideas
- opportunity to develop networks of support for resolving new problems
- powerful professional development for those involved
- mandatory participation and public scrutiny of one's own practices and judgments could encourage serious attention to the issue of quality
- being a useful supplement to other quality control procedures.

The disadvantages of assessor meetings include:

- meetings would need to be organised and facilitated (with attendant costs)
- would assessors come to meetings if they were voluntary? (or alternatively, what sanctions could be used to encourage participation?)
- there would need to be some fulltime officers to organise the process

(though the costs for this would probably be less than for moderators or panels)

- there are substantial logistic obstacles to covering all assessors
- without a formal approval process, there would be no guarantee of quality outcomes
- they would not satisfy the need for quality control in a high stakes situation.

### **Assessor partnerships**

Assessor partnerships simply involve sharing of assessments within a small group of assessors, maybe between just two assessors. This is a reduced form of assessor meeting, one where the public scrutiny is much reduced and the focus is on mutual assistance and confirmation.

Partnerships clearly have the potential to be of some assistance and are preferable to no action at all. However, external management is problematic. They must be self-managed. But this means that some will not treat such partnerships seriously and they may well have little effect. Also, some will treat the process as an imposition and resist or subvert it. Requiring sign-off means that many would treat it as a bureaucratic procedure rather than something of real benefit.

Assessor partnerships are small groups of assessors who:

- share and discuss each others' assessment approaches and materials
- provide mutual support for evaluating their assessment practices
- assist each other in resolving any puzzles or problems
- validate their assessment judgments in terms of the competency standards.

Partnerships could consist of only two assessors. The number of people in a partnership is limited only by the ability of the group to be self-managing and fully inclusive but would probably not involve more than five members.

Activities for assessor partnerships include:

- sharing ideas
- comparing notes
- resolving problems
- observing/commenting
- critiquing assessment materials
- reviewing competency judgments
- professional conversation.

Advantages of assessor partnerships include:

- they can be locally organised and do not need bureaucratic support
- there are few external costs apart from any promotional material
- partnerships can be promoted as providing mutual benefit to partners

## Quality management of school-based assessments

- participation is personally empowering, reducing uncertainties and enhancing assessment capabilities.

Difficulties with assessor partnerships include:

- it is possible that nothing may happen if it is voluntary
- participation depends on individual initiative and intrinsic motivation
- level of involvement depends on personal commitment
- successful partnerships require personal compatibility
- there may need to be some professional support and resources
- partners may simply reinforce each other's errors and misconceptions.

Assessor partnerships offer some important benefits but probably not by themselves. They can support other mechanisms of verification but do not, and are not intended to, provide a mechanism for verification themselves. The introduction of a requirement that they be part of the quality assurance requirements of registration would force them to be treated more seriously but would undercut their value as an agency for boot-strapping the quality of assessment practices through enlightened self-interest and mutual support.

*[This paper draws on material presented previously in Maxwell (2001) which provides a more extended treatment of the issues.]*

### References

Maxwell, G.S. (2001). *Moderation of assessments in vocational education and training*. Brisbane: Queensland Department of Employment and Training.  
<[www.trainandemploy.qld.gov.au/resources/about\\_us/pdf/moderation\\_report.pdf](http://www.trainandemploy.qld.gov.au/resources/about_us/pdf/moderation_report.pdf)>

Bransford J.D., Brown, A.N. and Cocking, R.R. (Eds). (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, DC: National Academy Press.

Linn, R.L. (1996). Linking assessments. In M.B. Kane & R Mitchell (Eds), *Implementing performance assessments: Promises, problems, and challenges* (pp. 91-105). Hillsdale, New Jersey: Lawrence Erlbaum.

Pellegrino, J.W., Chudowsky, N., and Glaser, R. (Eds). (2001). *Knowing What Students Know: The Science and Design of Educational Assessment*. Washington, DC: National Academy Press.