Are core learning outcomes 'standards'?

QUEENSLAND
SCHOOL
CURRICULUM
COUNCIL

Discussion Papers on Assessment and Reporting

Discussion Paper Number 1
Common and Different Features of Council and Board Approaches to Assessment and Reporting by Graham Maxwell, October 2001

Discussion Paper Number 2
Teacher Observation in Student Assessment by Graham Maxwell, October 2001

Discussion Paper Number 3
Are Core Learning Outcomes 'Standards'? by Graham Maxwell, March 2002

An earlier version of this paper was presented by Jim Tunstall and Graham Maxwell at the Australasian Curriculum, Assessment and Certification Authorities: *Bringing Assessment and Curriculum Issues Together*, July 2001.

In July 2002, the Queensland School Curriculum Council will amalgamate with the Queensland Board of Senior Secondary School Studies and the Tertiary Entrance Procedures Authority to form the Queensland Studies Authority.

PREFACE

This discussion paper is part of a series being published and disseminated by the Office of the Queensland School Curriculum Council. The purpose of this series is to encourage discussion on various issues concerning assessment and reporting.

This paper grew out of the realisation that discussions about 'standards' are often confusing because different people have different interpretations of the term. It seemed useful to clarify what those different interpretations might be. Also, there is continual comment, even heated argument, in educational circles about whether something more is needed in addition to a framework of core learning outcomes to allow reporting of student progress in terms of standards. This paper addresses that issue and suggests some ways of resolving it.

This discussion paper is not an official policy statement of the Council. Rather, it represents the views of the author, Dr Graham Maxwell, of the School of Education, The University of Queensland. As such, it offers a personal perspective on the issues. Dr Maxwell has been involved in research, writing and consultation on assessment for many years in Australia, USA and UK, ranging over all sectors and levels of education. He has taught courses and conducted workshops on assessment for pre-service and in-service teachers for 30 years. He has been involved in recent Council deliberations on assessment and reporting and written two other papers in this series. He was a member of the recent Taskforce on Assessment and Reporting of Education Queensland. He has been a member of the Moderation Committee of the Queensland Board of Senior Secondary School Studies for the past 13 years. He understands the local context well and speaks with authority on issues relating to assessment and reporting.

The audience for this discussion paper is professional educators, especially schoolteachers and administrators who must deal with assessment and reporting practice in classrooms and schools. The hope is that the discussion paper will serve as a basis for professional debate, professional development workshops and collaborative planning. It will also provide a useful reference point for discussions on standards under the new Queensland Studies Authority.

The Office of the Council would be pleased to have your reactions to and comments on the discussion paper, as well as any examples of ways in which the discussion paper may have helped to clarify or resolve any theoretical or practical problems you are facing in the implementation of Years 1 to 10 syllabuses. Your reactions, comments and examples would assist in deciding what further assistance might need to be provided on these issues.

JE Tunstall
Director
Queensland School Curriculum Council

March 2002

CONTENTS

INTRODUCTION

There are at least five different usages or meanings of the term 'standards' in relation to educational assessment and reporting. These are:

- Standards as moral or ethical imperatives (what someone should do)
- Standards as legal or regulatory requirements (what someone must do)
- Standards as quality benchmarks (expected practice or achievement)
- Standards as arbiters of performance quality (defining success or merit)
- Standards as learning milestones (progressive targets for student learning).

Whether core learning outcomes are 'standards' may depend on which of these meanings is intended. The following explanations help to clarify these meanings. Then the question can be asked whether core learning outcomes are 'standards' according to any of these meanings.

Here, reference to 'core learning outcomes' is in the context of current Queensland School Curriculum Council (QSCC) syllabuses for Years 1 to 10. In these syllabuses, core learning outcomes are represented as being 'essential' for all students. Other kinds of learning outcomes are included in these syllabuses, especially overall learning outcomes, key learning area outcomes and discretionary learning outcomes. Although the question about 'standards' could be asked with respect to each of these, the focus in this discussion paper is the core learning outcomes. Throughout, where principles may apply to any form of learning outcome, the word 'core' may be omitted. Even so, the issue addressed here is whether these core learning outcomes can be considered standards by at least one of the definitions of 'standards', not whether all possible forms of learning outcomes might be.

DIFFERENT MEANINGS FOR STANDARDS

Standards as moral or ethical imperatives (what someone should do)

Standards as moral or ethical imperatives offer guidelines for exemplary practice. They describe the way in which practice should be conducted, usually in the form of principles rather than operational procedures. There may or may not be consequences for any failure to implement or for any shortfall from appropriate implementation. For assessment practice, such standards are usually seen as desirable rather than essential.

Examples include:

- the recent *Guidelines for Exemplary Practice* section of the *Position and Guidelines on Assessment and Reporting* of the Queensland School Curriculum Council (QSCC)[1]
- the projected *Standards of Professional Practice for Accomplished Teaching in Australian Classrooms* of the Australian College of Education[2]
- the *Guidelines for Assessment Quality and Equity* of the Australian Curriculum, Assessment and Certification Authorities.[3]

---

[1] Queensland School Curriculum Council (QSCC) 2001, *Position and guidelines on assessment and reporting*, QSCC, Brisbane. [Replaced in 2002 by an updated document entitled *Position paper and guidelines on assessment and reporting*.]

[2] Australian College of Education 2000, *Standards of professional practice for accomplished teaching in Australian classrooms: A national discussion paper*, Deakin West, ACT. http://www.austcolled.com.au/Dispaper.pdf (accessed 24 August 2001).

[3] Australian Curriculum, Assessment and Certification Authorities 1999, *Guidelines for assessment quality and equity*. http://www.qbssss.edu.au/ACACA/guidelinesfor assessment quality and equity.html (accessed 26 October 2001).

Standards as legal or regulatory requirements (what someone must do)

Standards as legal or regulatory requirements form the basis of quality assurance approaches to system management. They usually focus on the formal fulfilment of particular processes or activities (and sometimes also on the provision of particular resources).

Examples include:

- the *Australian Quality Training Framework* (AQTF) for vocational education and training[4]
- the *Guidelines for Teacher Education Programs* of the Queensland Board of Teacher Registration[5]
- *ISO9000*.[6]

These kinds of standards assume that if the specified processes, practices and resources are 'in place', then desirable outcomes will be 'assured'. In other words, it is assumed that there is no need to check the quality of the resulting outcomes themselves. The focus is on quality of delivery or quality of practice. Typically, sanctions exist to enforce the standards. That is, these standards are not simply desirable; they are required.

Standards as quality benchmarks (expected practice or achievement)

Benchmarks indicate what ought to be the case, expectations of a particular level of quality in delivery or outcome. Benchmarks often have official endorsement and an implication of consequences if targets are not met. For example, sanctions may be imposed for failure to meet the benchmark or for falling seriously short. The *National Benchmarks for Literacy and Numeracy* in Years 3, 5 and 7 represent the assumptions underpinning the curriculum at these levels and therefore indicate the characteristics of literacy and numeracy all students ought to display as learning outcomes at these year levels.[7]

A looser form of this kind of standard is one set by particular people or interest groups according to their own view of what is desirable for educational practice or outcomes. Consequently, they apply this 'standard' to their own perceptions of the performance of students, schools or systems. The setting and interpretation of such a standard is typically implicit rather than explicit. Public debates about whether educational 'standards' of educational systems are satisfactory have this character. When people comment that system standards are low, they may be implying that the public benchmark is too low or that overall student performance is deficient (or both). Alternatively, there may be reference to the need to set 'high' standards or even 'world-class' standards in education and this has the implication of desired improvement in comparison with the past or in relation to other countries. Such

---

[4] Australian National Training Authority 2001, *Australian Quality Training Framework*, ANTA, Brisbane and Melbourne. http://www.anta.gov.au/aqtfImplementing.asp (accessed 24 August 2001). This framework replaced the previous Australian Recognition Framework (ARF).

[5] Board of Teacher Registration – Queensland 1999, *Guidelines on the acceptability of teacher education programs for teacher registration purposes*, BTR, Brisbane. http://www.btr.qld.edu.au/pdf/guide99.pdf (accessed 24 August 2001).

[6] International Organisation for Standardization (ISO) no date, *ISO 9000*, ISO, Geneva, Switzerland. http://www.iso.ch/iso/en/iso9000-14000/index.html (accessed 24 August 2001).

[7] Curriculum Corporation 2000, *Literacy and numeracy benchmarks*, Carlton South, Victoria. http://online.curriculum.edu.au/litbench/default.htm (accessed 24 August 2001). In this case, these are arbitrary standards, in the sense that they are socially decided not immutable, since they are designed to match a curriculum that is itself socially decided. They define what is needed to match the constructed curriculum. They are not 'natural' or 'functional' expectations.

low or that overall student performance is deficient (or both). Alternatively, there may be reference to the need to set 'high' standards or even 'world-class' standards in education and this has the implication of desired improvement in comparison with the past or in relation to other countries. Such standards are usually applied to systems as a whole rather than to individual students. The focus is usually on system accountability, particularly how teachers perform, rather than student accountability for their own learning.

Statements to the effect that a *school* has high standards imply that the school places demanding expectations on the students or that the aggregate performance of the students is high when compared to other schools (or both). Statements that a school has low standards imply that the school offers little challenge to students or that the aggregate performance of the students is low when compared to other schools (or both). In other words, these kinds of standards are about the relative level of expectation on students or their relative level of performance (or both).

These various interpretations of standards as quality benchmarks are what most people mean by the term 'standards' and what most public discussion about standards is about. However, there are other more neutral uses of the term 'standards'.

In both of the following uses of the term, a 'standard' is the referent for deciding whether student performance can be labelled as having certain characteristics. Whether the student performance is desirable, necessary or adequate are separate considerations (in the senses of the previous three uses of the term 'standards'), dependent on first identifying the student performance as being of a particular type or belonging to a particular category (such as a particular type or category of learning outcome). This is the focus of the following discussion.

Standards as arbiters of performance quality (defining success or merit)

Standards as arbiters of performance quality are referents that underlie human judgments of success or merit in performance. That is, they are the basis on which such judgments are made. Such a standard (or basis for the judgment) may be more or less implicit or explicit. Making the basis for the judgment (more) explicit facilitates sharing of understandings of the standard and developing of common interpretations of the standard. This helps to establish greater consistency of judgment among those making the judgment and to assist in making those judgments more transparent — this is the issue of assessment moderation and comparability, addressed in another discussion paper.

The point here is that these kinds of 'standards' are the basis on which the judgment is made about which label to apply to the performance. Examples are *competency standards* within the *Australian Quality Training Framework* (where there is one level of success for each unit of competency, that is, 'competent') and *differentiated grades* (for example, the five *Levels of Achievement* for subjects on the *Queensland Senior Certificate*). Assessment involves judging which of the designated 'standards' best represents the characteristics of the student's performance — in other words, what label to attach to the performance or what category to place it in.[8]

---

[8] The number of standards in each set depends on how the default standard is handled. In the case of competency standards, falling short of the competency standard ('competent') is typically represented as 'not yet competent' rather than 'failure', so there is really only one specified standard. An official result is recorded only when competence is demonstrated. In the case of graded assessments, sometimes the lowest grade is a default grade ('did not reach the next grade', or 'did not demonstrate a satisfactory level of performance' or 'failed') though sometimes the lowest grade is represented as requiring a minimum performance level and the default is not to receive any grade at all. The former is the case for Queensland Senior Certificate Levels of Achievement — that is, all students receive a result, no matter how poor the performance, provided they remain in attendance.

Standards as learning milestones (progressive targets for student learning)

Standards as learning milestones are the referents that underlie human judgments of progress in learning. Similar to standards as arbiters of performance quality, they are the basis on which judgments of progress are made. They also can be more or less implicit or explicit. They represent sequential targets for student learning and provide the point of reference for assessing what stage of development has been reached along a continuum of learning outcomes. A framework of learning outcomes is provided as a map against which progress can be charted. The aim of assessment is to chart the developmental progress of the individual learner in order to report that progress and/or to provide deliberate starting-points for further learning.[9]

In Australia, the various State and Territory syllabus frameworks of core learning outcomes in key learning areas purport to possess these characteristics.[10] In QSCC syllabuses, sequences of core learning outcomes are defined. The learning progress of the student, as revealed by demonstration of the core learning outcomes, can be recorded as a detailed profile across these sequences. Assessment is intended to provide students with opportunities to demonstrate which level on each sequence of core learning they have reached. In other words, assessment should allow determination of which core learning outcomes have been demonstrated, not simply whether the student has 'passed' or 'failed' at a particular level.[11]

DELIVERY STANDARDS VERSUS OUTCOME STANDARDS

A distinction can be made between standards relating to delivery and standards relating to outcomes. The first two categories of meanings for standards discussed above emphasise delivery standards, the last two categories emphasise outcome standards, and the middle category covers either or both. Delivery is concerned with activities or processes that are considered to support student learning. Outcomes are concerned with learning targets to which the delivery is directed.

The argument for defining curriculum in terms of sequences of core learning outcomes is that the focus is then on ensuring that all students are assisted to attain desirable (core) knowledge, capabilities and attributes. The identified learning outcomes exert a 'pull' on teaching and learning processes. They also serve as the basis on which judgments about learning progress are made. Consequently, in this approach to curriculum, planning for teaching and learning is coincident with planning for assessment and reporting. A framework of core learning outcomes functions as both a curriculum framework and an assessment framework. In this sense, the core learning outcomes define the outcome standards — as targets for student learning and as referents for whether that learning is occurring.

There are two requirements for learning outcomes to function as satisfactory outcome standards in this sense, that is, as satisfactory sequential learning targets. First, the framework of learning outcomes needs to be *defensible* in terms of its content, structure and sequencing. This is a question of the appropriateness of the overall curriculum design. Second, the stated learning outcomes need to be *adequate* for making appropriate distinctions between students who have and have not demonstrated

---

[9] For a comparison of the 'merit' versus 'milestones' approaches to assessment, see the Assessment and Reporting Discussion Paper *Common and Different Features of Council and Board Approaches to Assessment and Reporting*. Here, the focus is on the 'milestones' approach of QSCC syllabuses.

[10] The various State and Territory curriculum boards are accessible through the QSCC website (as education links) or directly at http://www.qscc.qld.edu.au/education_links/index.html

[11] The thinking associated with this approach is given in Queensland School Curriculum Council (QSCC) 2001, *Position and guidelines on assessment and reporting*, QSCC, Brisbane.

those learning outcomes. While both defensibility and adequacy are important considerations, this paper considers only the latter.

WHAT DOES IT MEAN TO USE LEARNING OUTCOMES FOR MAKING ASSESSMENT JUDGMENTS?

In some places, assessment against core learning outcomes (or 'milestones') has been defined as 'standards-based assessment'.[12] However, so too has assessment against merit standards (or 'grades'). In Queensland, there has been reluctance to label assessment against core learning outcomes for Years 1 to 10 as standards-based assessment because of differences from and possible confusion with the rather different form of standards-based assessment adopted in Years 11 and 12 — merit standards for 'Levels of Achievement'. Because of its longer history of use, the term 'standards' carries strong 'grade' connotations in these circumstances. This has led to the strong sentiment in some quarters concerning Queensland syllabuses for Years 1 to 10 that 'outcomes are not meant to be standards'. Such a statement needs to be seen in its historical context. From a broader perspective, both types of assessment are standards-based and share many other features in common despite their important differences.[13]

Standards for criteria and criteria for standards

Following Sadler (1985, 1986, 1987), the Queensland approach to assessment in Years 11 and 12 is typically referred to as 'criteria-and-standards-based assessment' (though sometimes 'criteria-based assessment' or 'standards-referenced assessment'). Essentially, this involves the use of a rubric (or schema) to represent the conjunction of different criteria (or performance dimensions) with different merit standards.

As discussed by Sadler, a hierarchy of criteria can be recognised with standards on the criteria at one level being characterised by sub-criteria at the next (lower) level. That is, standards elaborate each of the criteria but (sub)criteria also elaborate each of the standards. This hierarchical structure is shown in appendix 1 (from Sadler 1986).

Similarly, if core learning outcomes are considered as standards, then each standard can be characterised in terms of relevant (sub)criteria that make its meaning clearer. Again, though, because of the strong association in Queensland of the term 'criteria' with rubric dimensions (from Years 11 and 12 syllabuses), it has been more common to refer to such (sub)criteria (in Years 1 to 10 syllabuses) as 'elaborations' (in common with other States and Territories). Whatever terminology is used, it should be clear that it is appropriate to explicate the meaning of a learning outcome in terms of more detailed characteristics. These characteristics can be further explicated in terms of their sub-characteristics, and so on. It is a matter for pragmatic decision how much detail is needed and how many hierarchical stages (or further levels of explanation of meanings) are needed.

---

[12] See, for example, Lambert (no date): *Standards referenced assessment in primary schools*.
http://www.bosnsw-k6.nsw.edu.au/parents/k6standards_assess.html (accessed 02/08/2001).
In fact, a more generic term for the assessment approach espoused by Lambert would be 'outcomes-referenced assessment'. Unfortunately, though, in equating outcomes with standards, Lambert confuses milestones and benchmarks — reference points for determining what progress in learning has occurred versus expectations of what students ought to demonstrate by particular stages or year levels of education. QSCC syllabuses do not make very strong statements about the year levels by which core learning outcomes *ought* to be demonstrated, except that most students should eventually reach level 6.

[13] See footnote 10.

One conceptual distinction should be noted. In an assessment rubric or schema, the criteria are seen as defining generic dimensions on which several merit standards can be represented (usually but not necessarily with the same number of merit standards on each of the criteria). The criteria are typically 'abilities': process or skill categories. For example, in the Queensland Senior Physical Education syllabus, the three overall criteria for the subject are:

- *acquiring* (referring to abilities such as gathering, recalling and comprehending information in, about and through physical activity)
- *applying* (referring to abilities such as interpreting, analysing and manipulating information in, about and through physical activity)
- *evaluating* (referring to abilities such as hypothesising, justifying and appraising information in, about and through physical activity).

Although standards in a rubric are merit distinctions rather than developmental stages, each of the standards statements taken by themselves can be considered as a learning outcome statement. For example, in the Queensland Senior Physical Education syllabus referred to above, the following extracts of exit standards appear:[14]

- Gathers information relating to physical activities, showing the use of categorisation and reasoning (part of standard B for acquiring)
- Applies physical responses in simple performance environments (part of standard D for applying)
- Produces credible and convincingly justified responses evaluating physical activities (part of standard A for evaluating).[15]

These standards statements are fairly typical of Queensland Senior syllabuses.[16] They do not look very different, lexically and grammatically, from statements of core learning outcomes in QSCC syllabuses. This point is revisited later.

Standards are necessarily fuzzy

No matter what degree of explanation is used to define a standard, the standard is necessarily 'fuzzy'. That is, standards cannot be expressed with such precision that the standard is completely unambiguous and will be apprehended identically by all assessors. Statements of standards require interpretation and this can differ according to the meanings ascribed to the words. These meanings need to be generated within the community of users. There are three important ways in which such meanings are generated: through relationships among and explanations of the learning outcomes in the syllabus materials; through professional development workshops for teachers; and through moderation meetings among teachers.[17] The first of these — written documentation — is relatively weak, by itself, for producing common understandings. Also needed are meetings among teachers to discuss and verify their interpretations and understandings of syllabus statements of learning outcomes and to share examples of student performances that they judge to be demonstrating those learning outcomes (through professional development workshops or moderation meetings or both).

---

[14] These are only part of the relevant standards statements.

[15] For each of these criteria there are five standards (A to E) that map through a system of trade-offs onto the five overall Levels of Achievement.

[16] As would be expected, recent Senior syllabuses show more extended elaboration of the standards. However, this does not detract from the point being made here, since simpler statements have sufficed in the past. In fact, this provides support for the value in adopting an action learning cycle (see later).

[17] Moderation is here understood to be a social process involving comparison of assessor judgments, not a statistical process involving scaling of 'marks' or 'scores'.

Some people (e.g., Cresswell, in Goldstein and Heath 2000) claim that it is impossible to state standards with sufficient clarity and that attempts to do so are therefore doomed to failure. Ultimately, however, the question is what degree of ambiguity of interpretation of the learning outcomes is tolerable for the context in which they are used. This can be translated into: what are the characteristics of a well-stated standard; and what supporting processes of professional development and moderation are needed to assist in delivering an acceptable degree of commonality of meaning? Defining what is an acceptable degree of commonality (or tolerable degree of ambiguity) requires widespread debate and agreement. The issue can only be resolved by balancing costs and consequences. What is clear is that common interpretations of learning outcomes cannot be guaranteed through printed documents alone.

What characteristics of a learning outcome make it a well-stated standard?

Wiggins (1998) argues convincingly that well-stated learning expectations (and assessment specifications) should include three components:

- a doing component (what activity or process is involved?)
- a knowledge component (what is the intellectual content?)
- the criteria for satisfactory performance (what distinguishes those who 'understand' from those who do not).[18]

Many statements of learning outcomes are relatively strong on the first two components but rather weak on the third.[19] This point is explored later in relation to particular core learning outcomes. However, here, there are some additional issues worth considering. Typically, statements of learning outcomes are fairly brief statements. In themselves, they are unlikely to provide much detail concerning the nature of the activity or process and the knowledge involved. They are likely to be indicative of the type of activity/process involved and the type of knowledge involved rather than to provide specific details. In this sense they encapsulate and signify the learning outcome rather than offer specificity and completeness.

Core learning outcomes are not, however, interpreted in isolation. First, they exist in relation to the whole framework of core learning outcomes (within and across key learning areas) and in relation to general principles, perspectives and priorities (within and across key learning areas). Second, the core learning outcomes can be supplemented by additional details in the form of elaborations. Third, core learning outcomes can be exemplified by shared examples of assessment materials illustrating student demonstrations of particular learning outcomes. The importance of exemplars for creating common understandings of standards is now well accepted.

QSCC syllabuses and syllabus materials support the interpretation of learning outcomes in all three ways. First, the learning outcomes are structured through strands, key concepts and/or processes, and sequences (across levels). Also, various principles, perspectives and priorities infuse the learning outcomes. These include general principles and emphases of the eight Key Learning Areas, the eight overall attributes of lifelong learners, and several cross-curricular priorities (literacy, numeracy, lifeskills and futures perspective). Some syllabus materials also make explicit connection to related aspects of other Key Learning Areas. The characteristics of the learning outcomes therefore need to be interpreted within the overall framework and philosophy of the curriculum.

---

[18] For Wiggins, understanding encompasses application and skill (doing) as well as apprehension (knowing).

[19] This is not necessarily the case for all statements of standards. For example, the examples given earlier of stated standards in QBSSSS Physical Education emphasise the activity (or process) but are weak on content. Beyond this, they are silent on the characteristics of satisfactory performance for that level, as also are QSCC statements of core learning outcomes.

Second, the syllabus materials provide elaborations of the learning outcomes to support their interpretation. These include elaborations on both content and processes — that is, on both the knowing and the doing aspects of the learning outcomes. While these elaborations should provide some assistance in encouraging consistency of interpretation of learning outcomes, they are unlikely to be sufficient. Other more direct means of developing consistency of interpretation are likely to be needed. The need for this will, however, have to arise out of the experience of implementation since there is no current mechanism for going further. Unless there are consequences from different interpretations and implementations by teachers and schools, this situation is unlikely to change. More use would need to be made of teacher-based assessment and reporting — for example, contributing to school level and system level accountability — for this to happen.

Lastly, QSCC has been exploring the use of annotated work samples in assessment and reporting. Annotated work samples could highlight, for particular examples of student work, the performance characteristics that contribute to demonstration of particular learning outcomes. Such annotated work samples would be illustrative rather than exhaustive, that is, would not cover all core learning outcomes. However, they could provide exemplars for some core learning outcomes and constitute a basis on which further exemplification of core learning outcomes could be built. As is always the case with exemplars, it would be important to exemplify the learning outcomes in a variety of ways and to update these regularly to prevent ossification.

The role of judgment

An important issue to consider is the role of human judgment in the use of a standard. Not only must the meaning of the standard itself be constructed in each teacher's mind (involving judgment) but evidence must also be interpreted by comparison with the standard (also involving judgment). Whether the results of these processes have a high degree of consistency depends on the supports that are put in place for creating convergence among teachers. QBSSSS research has shown that where assessors are properly trained and monitored a high degree of consistency can be attained.[20] Where training and moderation is weak or non-existent, it should be expected that consistency will not be high. The situation for QSCC syllabuses is unknown.

The QSCC position on assessment for Years 1 to 10 syllabuses (QSCC 2001) is that it should:

- focus on the core learning outcomes
- be continuous and classroom-based
- involve collection of evidence on the demonstration of the core learning outcomes
- allow for multiple opportunities to demonstrate the core learning outcomes
- take account of a variety of contexts (variations of content, style and setting)
- involve on-balance professional judgments (interpretation of the evidence)
- produce a profile of demonstrated learning outcomes (across levels)
- serve both formative (teaching) and summative (reporting) functions.

This type of assessment is different in important respects from previous approaches to assessment. Such assessment is evidential, contextual, continuous, judgment-based and standards-referenced. Crucially, too, such assessment does not need to involve standardised tasks. Rather, assessment can be individualised without compromising the need for fairness and equity. *In fact, the situation is quite the opposite — the greatest degree of fairness and equity can be attained by ensuring that all students have appropriate opportunities to demonstrate the characteristics of their learning.* Different evidence may be appropriate for different students. The important requirements are that the evidence be relevant to the learning outcome(s) and the judgment defensible in terms of the requirements of the learning

---

[20] See reports by QBSSSS on the moderation process for Senior subjects and on marking the Writing Task which forms part of the Queensland Core Skills (QCS) Test.

outcome(s). This does not need to involve identical evidence, or even identical interpretation of the evidence. It is sufficient to strive for equivalence, for example, agreement that two different collections of evidence do (or do not) reveal equivalent demonstrations of the same learning outcome(s). This philosophy also applies in QBSSSS approaches to assessment.

SHOULD MERIT STANDARDS BE ATTACHED TO LEARNING OUTCOMES AS WELL?

It has been argued here that core learning outcomes, as stated in QSCC syllabuses, can be considered as representing performance standards in the sense of providing a descriptive baseline for judging whether a student has demonstrated each core learning outcome. They possess two of the three components that are desirable for well-written performance standards, that is, reference to 'content' and 'activity'. They do not possess a third desirable component, that is, 'criteria for satisfactory performance'. In this respect, they are not different from other examples of performance standards in practical use, for example, performance standards for QBSSSS Exit Levels of Achievement. In fact, performance standards for QBSSSS Exit Levels of Achievement typically lack reference to 'content' as well.

It may be argued that performance standards for QBSSSS Levels of Achievement as expressed in subject syllabuses are only part of the picture and that these are elaborated further in school assessment programs, in terms of all three desirable components of performance standards. However, the same could be claimed for QSCC core learning outcomes. What sets the two approaches apart is not the formal expression of a performance standard but the existence of a moderation system for assessments of QBSSSS Exit Levels of Achievement and the absence of a moderation system for QSCC core learning outcomes. Some form of moderation is clearly needed if performance standards of any kind are to be interpreted and enacted in more or less common ways by teachers. The burden of common meaning cannot be carried by written statements alone.

All of this, while important, may miss the point. Perhaps the real issue in the debate about performance standards is whether standards as learning milestones (progressive targets for student learning, as expressed by core learning outcomes) are sufficient. In other words, should core learning outcomes be supplemented or replaced by one or other or both of the other kinds of performance standards, that is, standards as arbiters of performance quality (such as merit grades or levels of achievement) or standards as quality benchmarks (such as typical targets for particular year levels)?

The answer to this question may differ according to the perceived purpose of assessments — that is, 'sufficient for what use?' Formative purposes of assessment, in the classroom, are served better by more detail. For strategic teaching, attention to even more detail than that provided by core learning outcomes is probably necessary. Nevertheless, core learning outcomes can function effectively as identifiable and manageable targets for each student's next stage of learning and encourage a tailored approach to addressing each student's developmental needs. On the other hand, summative purposes of assessment may be served better by less detail, but the level of detail may need to differ according to the focus of reporting —individual students, schools or systems. Even so, core learning outcomes can function as the recorded data from which different types of reports can be generated for different summative purposes.

It is possible to generate grades from a database of core learning outcomes by specifying combinatorial rules (using formulas or trade-offs). Some schools have already done this for purposes of reporting an exit grade at the end of Year 10. This allows a profile on several sequences of learning outcomes to be collapsed onto a single dimension for each key learning area (or possibly each strand), providing relative comparisons of overall student performance. This may satisfy a felt need for comparisons of this type at particular junctures of schooling. However, before rushing to do this, there are some issues that ought to be considered.

First, it is somewhat against the philosophy and design of an outcome-based curriculum, and thus of QSCC syllabuses, to introduce grades into the assessment. The emphasis is instead on recording and reporting a profile of what each student knows and can do. The framework of core learning outcomes facilitates judgments, on the basis of collated evidence, of the level the student has reached on each sequence of core learning outcomes. This represents *how far* the student has progressed rather than *how well* the student has performed.[21] *How far* carries positive connotations of development and potential improvement, whereas *how well* carries connotations of competition and potential failure.[22] In other words, there are some possible benefits in the outcome-based approach that should not be lightly thrown away.[23]

Second, unless grades are defined centrally, any formulation adopted locally has limited meaning and value. Even if defined centrally, grades would still have limited interpretability and usefulness unless the grades at one juncture of schooling could be connected to — that is, mapped onto — the grades at another juncture. This would be a difficult though not impossible proposition to implement. One advantage of the framework of core learning outcomes is that there is a single framework applicable across Years 1 to 10. Ultimately, when fully operational, this could provide an efficient reference point for all assessment across those years, with student records and portfolios being cumulative across years, able to be handed from one teacher to the next and being interpretable by all teachers in the system. Student development could be systematically and coherently charted over time. Through familiarity of use by all teachers, the framework of core learning outcomes could provide a standard reference point for all reporting. The same process of judgment is involved at all year levels against the same familiar framework. This judgment involves deciding on the basis of the available evidence what level the student has reached on each of the (relevant) sequences of core learning outcomes. This degree of coherence across years holds more potential for tracking and comparing student development than arbitrary grades at particular junctures of schooling.

Of course, what may be worth exploring is whether profiles (of core learning outcomes) and grades (for more global outcomes) might both be reported at particular junctures of schooling. This would serve two purposes (structural coherence as well as summary comparison). Any acceptable process of synthesising demonstrated core learning outcomes into grades would require transparency and simplicity. Meaningful representations of the standards for the grades would need to be developed and these would need to relate to relevant core learning outcomes. This would go much beyond the present situation by defining performance expectations (quality benchmarks) for particular junctures of schooling.

An advantage of grades is that they offer a combination of summarisation, comparison and benchmarking — providing a simple and direct representation of how well the student has performed, how the performance compares with that of other students, and how well the performance compares with desired learning targets (as expressed by the standards). On the other hand, a disadvantage of grades is that they are not connected to a developmental continuum, therefore allowing only vague interpretations of progress over time. That is, the specifics of the grade standards, the meanings of the grade labels, typically differ in idiosyncratic ways across different contexts (especially, different years). Tracking progress over time therefore becomes a complex (and somewhat ambiguous) process, requiring interpretive information that is usually neither explicit nor accessible. At face value, a student may appear to make no progress over time, for example, if the student always receives a C-grade. As well, this can lead to reification of the student's 'ability' — characterisation as a 'C-grade

---

[21] *How far*, in this case, also implies increasing depth and sophistication of knowing and doing.

[22] The standard *how well* approach to assessment creates winners and losers. Differential standards are pitched to the likely range of student performance and adjusted on the basis of empirical experience. The standard psychometric model underpinning this approach assumes there is a (natural) distribution of ability in the student population and that assessment is directed at representing that distribution.

[23] These benefits are explored in detail in the QSCC Outcomes Information Package: http://www.qscc.qld.edu.au/p-10_framework/index.html#outcomes.

student'. Another feature of grades is that slower progress tends to be characterised negatively ('failing'; 'limited') rather than positively (in terms of the actual knowledge and skill that has been demonstrated). Grades create 'winners' and 'losers'.

While the framework of core learning outcomes provides a coherent structure for charting student progress over time, the size of the jump between levels is perceived by some to be rather large. That is, given six levels along a continuum across 10 years, there is an average time span between levels of about 20 months. A student may appear to be marking time between levels, that is, making no progress for a long time, in some cases for even longer than 20 months. It must be remembered, however, that there are other checkers on the board and they may not all move at the same rate. That is, progress on different sequences of core learning outcomes is likely to be uneven. Profiles typically will show that the student has demonstrated core learning outcomes at various levels. Comparison of profiles from year to year can be expected to show that the student has made progress on some core learning outcomes but not others.[24]

If desired, interim progress on core learning outcomes can be captured in various ways without resorting to grades. One possibility is to introduce interim steps between successive core learning outcomes. A three-point scale is suggested in some quarters, for example, 'working towards', 'consolidating', 'demonstrated'. This would at least preserve the coherence of the framework of core learning outcomes. However, although this approach has been adopted elsewhere in Australia, it lacks rigour. Without the expression of explicit standards for these scale points, their meaning remains extremely vague and their application will be inconsistent. The benefit of such graduated steps between core learning outcomes is unclear. Furthermore, their use increases (triples) the number of decisions to be made about progress on the core learning outcomes.[25]

Another possibility for capturing interim progress is a teacher commentary on the available evidence about the progress the student is making towards the core learning outcomes that have not yet been demonstrated completely. Whether these details are recorded formally or generated verbally according to need (for example, during discussions with parents), they should always, in principle, be retrievable from a portfolio of student work.

The adoption of grades would move in the opposite direction, with larger quantum jumps between the reporting steps. Grades necessarily represent broader ranges of performance, functioning as summarising devices by classifying students into a small number of global categories. Students in the middle categories display different mixes of performance characteristics. The loss of detail is balanced by the gain in simple summarisation.

A different way of summarising would be through the so-called level statements (for strands). This would maintain the integrity of the existing outcomes structure but allow more 'chunking' of the information on performance. The level statements purport to summarise the core learning outcomes for a level on a strand. However, these are currently written with fewer characteristics of 'standards' than the core learning outcomes — essentially summarising the content of the relevant core learning outcomes at that level. For this reason the *QSCC Position and Guidelines on Assessment and*

---

[24] This may depend on opportunities for learning and opportunities to demonstrate the core learning outcomes and these opportunities depend on the school's curriculum and assessment plan.

[25] A different amount of detail may be kept at different stages of assessment. Thus, at the classroom level, a student portfolio may contain essentially all the student's work (raw data) including such things as annotated work samples, observation records, running records, progress summaries and feedback comments. At the school level, a periodically updated record (table) may be kept of demonstrated core learning outcomes. What level of detail is desirable for reporting to parents is a matter to be negotiated between the parents and the school. A distinction needs to be made between what is noted (a classroom level), what is recorded (at school level) and what is reported (to parents). A distinction should also be made between what is reported to parents on a formal report card and what might be discussed with parents drawing on more detailed information from student portfolios and teacher records.

Reporting recommends against the use of trade-offs in reporting progress in terms of the level statements. Rather, the guidelines suggest that demonstration of a level statement should require demonstration of all the core learning outcomes covered by the level statement. In that case, reporting on the level statements adds little to reporting on the core learning outcomes separately — it merely adds emphasis. Even so, this is likely to be unworkable, since it assumes that progress will be consistent across the core learning outcomes within each strand. Conflicts of interpretation will arise if this is not the case. If level statements are to be used meaningfully for reporting progress, they may need to be redrafted to satisfy the requirements for a well-written standard more closely — that is, content, activity and criteria — and allowance made for on-balance judgments (involving trade-offs) concerning whether they have been demonstrated or not.

On the issue of whether there also should be quality benchmarks associated with core learning outcomes, to define them as standards in this sense, it is clear that this could be done, as has been the case in some other Australian states. However, this requires some empirical validation of the sequences of core learning outcomes and some knowledge of what progress students typically make as well as what progress we should preferably desire. In the meantime, it would be helpful to begin building the database of normative and comparative data to allow individual student progress to be represented in relation to other students in the school, other schools and the State.

WHAT SUPPORT IS CURRENTLY PROVIDED FOR CORE LEARNING OUTCOMES AS PERFORMANCE STANDARDS?

The expression and interpretation of learning outcomes as performance standards should be seen as an ongoing process, akin to an action learning cycle. This action learning cycle might be represented as follows:

- expressions of learning outcomes are constructed (and refined) within the process of syllabus development
- further refinement occurs within the trial process leading to reformulated expressions of the learning outcomes
- supplementary materials are developed to assist in the interpretation and use of the syllabus and to clarify the meaning of the learning outcomes
- the syllabus is released and used by teachers who construct their own meanings and enact their own interpretations of the learning outcomes
- sharing of these meanings and interpretations occurs among teachers (opportunistically through teacher contacts and networks or deliberately through professional development workshops and moderation meetings)
- teachers' experience with the syllabus provides a basis for future revisions of the syllabus during which further refinements of the expression of the learning outcomes may occur.

For QSCC syllabuses, the first four of these stages have been occurring. There are currently no plans for the last two stages. Because assessment was not seen as part of QSCC's brief, the process of refinement of the syllabus and supplementary materials has not included any substantial evaluation of the usefulness of the learning outcomes as performance standards. However, the question is not whether they are completely satisfactory but whether, taken in context, they provide an adequate starting-point.

It is instructive to look at some examples of learning outcomes and support materials. In doing so, we point to some of their characteristics without drawing strong conclusions about their adequacy. As in all educational practice, it is more important to ask 'what should we do next?' rather than 'why didn't we do better?'. In general, the examples indicate substantial progress in a 'learn as we go' environment. The attached examples are drawn from Science and Studies of Society and Environment.

Appendixes 2 and 3 are extracts from the Years 1 to 10 Science syllabus and sourcebook guidelines for the Life and Living Strand, which has three concept-related sequences of learning outcomes across the six 'levels'. Three statements from the Sourcebook are listed to emphasise the context in which the learning outcomes are to be interpreted. The sequencing is seen to involve development of increasing complexity (cognitive and conceptual). The sequencing also encourages interpretation of each learning outcome in relation to the whole. The learning outcomes are seen to include a *doing* component and a *content* component. In this sense, these learning outcomes cover two of the components of well-stated learning expectations noted earlier (from Wiggins 1998). However, the third component — criteria for satisfactory performance, that is, distinguishing those who 'can' from those who 'cannot' — is missing. Elaborations of the learning outcomes in the Science: Years 1 to 10 Sourcebook Guidelines provide substantial additional detail on the content component but no further comment on the activity/process component. Criteria for successful performance are not identified and will need to arise out of practice in implementation.

Appendixes 4 and 5 show extracts from the Years 1 to 10 Studies of Society and Environment syllabus and sourcebook guidelines showing the core learning outcomes for all strands, conceptual organisers and processes. These learning outcomes also include doing and content components. In this case the elaborations provide additional details on both these components. Again there are no criteria for successful performance.

The more recent Studies of Society and Environment syllabus shows greater sophistication in the structuring and expression of learning outcomes than the earlier Science syllabus. However, the absence of criteria describing the characteristics of successful performance means that the learning outcomes in both syllabuses are limited in their capacity to function as standards. Nevertheless, the basis is there and other stages of the action learning cycle can assist in adding 'understandings in practice' about the third component.

The absence of characteristics for distinguishing between successful and unsuccessful demonstration of core learning outcomes probably is not fatal for their use as performance standards. After all, as has been shown, QBSSSS statements of standards similarly lack precision.

SOME DIRECTIONS FOR FUTURE DEVELOPMENT

The focus in QSCC syllabus development was mainly on the 'curriculum' aspects of syllabuses, that is, the teaching and learning. Assessment was defined as the responsibility of the school authorities. Over time there has begun to develop a greater appreciation of the interrelatedness of curriculum and assessment, with some re-visitation of assessment issues — as evidenced in the recent *Position and Guidelines on Assessment and Reporting* (QSCC 2001). Perhaps this will extend further in the future.

Some thoughts for consideration in the future include the following:

- For learning outcomes to serve as standards, consideration needs to be given to the way they are expressed and illustrated. We should look at the precision of the language used (especially the doing component), the need for elaborations to cover both the content and the doing components, the expression of criteria for satisfactory performance, and the provision of exemplars to illustrate typically satisfactory performance through concrete representations.

- Assessment needs to be considered as an integral aspect of the development of syllabuses that are built around learning outcomes. Learning outcomes serve both a curriculum role (structuring and focusing for teaching and learning) and an assessment role (representing the basis on which learning is recognised and reported).

- Typical statements of merit standards (as, for example, with Queensland Year 12 Exit Levels of Achievement) are not expressed as clearly as many milestone standards (as, for example, with Queensland Years 1 to 10 core learning outcomes). The need for improvement is not one-sided.

- For common understandings of learning outcomes to develop, teachers need assistance in interpreting and implementing existing syllabus frameworks. Consideration needs to be given among all the stakeholders about how this could best be done. Implementation of the new syllabuses requires reconstruction of practice. Previous research on educational innovations suggests that this is unlikely to occur systematically without appropriate support systems.

- Common interpretations of learning outcomes also require some form of sharing among teachers. Such understandings develop best when teachers share examples of student performance and debate their separate judgments of whether they represent a demonstration of particular learning outcomes. This could involve some form of moderation among teachers, one based on professional consultation, not necessarily one involving external accountability. Such professional consultation and sharing generate very powerful professional development.

The answer to the question whether core learning outcomes are standards is 'potentially so'. Some other stages of the action learning cycle need to occur for this realisation of the potential. The advantage of seeing core learning outcomes as standards — milestone standards, not merit standards or benchmark standards — is that it helps us focus attention on sharpening our articulation of the core learning outcomes, providing a firmer basis on which comparable judgments can be made by teachers and providing a stronger foundation for using teacher judgments of student demonstration of core learning outcomes as the basis for school and system reporting on educational outcomes. Whether merit standards and/or benchmark standards are desirable as well as, or instead of, milestone standards is an issue needing further exploration.

REFERENCES

Goldstein, H. & Heath, A. 2000, 'Educational standards', *Proceedings of the British Academy*, 102, 69-158.

Maxwell, G.S. 2001, *Common and different features of Council and Board approaches to assessment and reporting*, Queensland School Curriculum Council, Brisbane.

Queensland School Curriculum Council (QSCC) 2001, *Position and guidelines on assessment and reporting*, QSCC, Brisbane. [Replaced in 2002 by an updated document entitled *Position paper and guidelines for assessment and reporting*.]

Sadler, D.R. 1985, 'The functions and origins of evaluative criteria', *Education Theory*, 35, 285–297.

Sadler, D.R. 1986, *General principles for organising criteria,* (Discussion Paper 9), Queensland Board of Senior Secondary School Studies, Brisbane.

Sadler, D.R. 1987, 'Specifying and promulgating achievement standards', *Oxford Review of Education*, 13, 191–209.

Wiggins, G. 1998, *Educative assessment: Designing assessments to inform and improve student performance*, Jossey-Bass, San Francisco.