



## *ASF Working Paper 1*

# **Aims and outcomes of the first year's work of the project**

## ***ASF Project Working Paper 1***

### **Contents**

#### 1. Introduction

#### **Part 1**

#### 2. Background, aims and procedures of the project

##### 2.1 Background

##### 2.2 The project goals

##### 2.2 Procedures

#### 3. Outcomes of the first year's work

##### 3.1 Clarification of terms

##### 3.2 Developing principles of summative assessment by teachers

##### 3.3 Describing and analysing methods of summative assessment by teachers

#### **Part 2**

#### 4. Definition of summative assessment by teachers

#### 5. Principles of summative assessment by teachers

##### 5.1 Overall goals

##### 5.2 Principles

##### 5.3 Managing assessment by teachers in a summative assessment system

#### 6. Model for describing and evaluating methods summative assessment by teachers

##### 6.1 Components of summative assessment methods using TA

##### 6.2 Representing the dynamic relationships among the components

#### **Part 3**

#### 7. Reflection, synthesis and implications

##### 7.1 The evidence and policy context

##### 7.2 Implications

#### **Appendices**

Appendix 1 Key points from practice in different countries

Appendix 2 Key points from systematic reviews of research related to teachers' summative assessment

# 1 Introduction

The Assessment Systems for the Future (ASF) project is a project of the Assessment Reform Group ARG. Its focus is summative assessment and the role that assessment by teachers can take in it. In Part 1 (sections 2 and 3) of this paper we explain the reasons for the focus on summative assessment by teachers, the project procedures and the outcomes of the first year's work. In the first year of our work we conducted three seminars, of which the outcomes are summarised in Section 3 (more detailed reports of each seminar and the papers presented are available on the Assessment Reform Group website<sup>1</sup>.) The aim of these seminars has been to bring together comprehensive analysis, critical review of evidence, reflections both practical and theoretical on issues of systemic dynamics and strategies for change, and to illuminate our overview by looking at different systems within and beyond the UK. Such a synthesis will inevitably demonstrate that there are no easy answers, but it is hoped that the outcomes of the project can offer guidance both on possible paths to change which might well be pursued, and to those which ought to be avoided.

Part 2 sets out the thinking that has emerged from the first year's work. Section 4 gives detail of the definition of summative assessment by teachers that emerged from discussions in the project. In section 5 we identify some principles for guiding the development of procedures and systems for summative assessment that makes use of teachers' judgements. We also touch upon some practical matters of managing such systems, although these matters will feature more prominently in later discussions within the project as we move towards specific recommendations. Section 6 sets out our thinking about the components of teachers' summative assessment procedures and an approach to analysing the relationships among these components.

Part 3 brings together the outcomes of consultation on the implications of our work with policy-makers and practitioners from a range of institutions and organisations to arrive at proposals. These will be further refined in the course of the seminars and conferences to be held in the second year of the project.

---

<sup>1</sup> The ARG website [www.assessment-reform-group.org](http://www.assessment-reform-group.org)

## **Part 1**

### **2. Background, aims and procedures of the project**

#### **2.1 Background**

The previous work of the ARG has shown that assessment can be used to help students' learning<sup>2</sup> but also that assessment systems that rely on tests can have a narrowing impact on teaching and on the curriculum and a negative impact on students' motivation for learning<sup>3</sup>. Further, it has become clear from research evidence that teachers' ability to take advantage of the benefits of using assessment formatively, to help learning, is being impeded by the over-emphasis in some assessment systems on external tests and examinations, particularly when the results are used for high stakes purposes.

Summative assessment has a key role in education and has a range of uses. The challenge is to conduct it in such a way that minimises the negative effects that have been documented. Assessment by teachers has the potential for providing summative information about students' achievement across the full range of activities and goals without having the negative effects on teaching and the curriculum that are associated with tests. However, although assessment by teachers is used as the main source of information in some national and state assessment systems, in other countries it has the image of being unreliable and subject to bias. The implicit assumption that external tests and examinations are reliable measures has to be challenged. Published research on this matter is limited but what there is indicates that an unacceptable proportion of students are placed in the wrong levels or grades.

The way forward is not to replace tests entirely by teachers' assessment but to consider what would constitute an optimum system for summative assessment. In this context 'assessment system' refers to far more than the procedures for collecting data about students' performance. Its meaning extends to the way in which the data are communicated, reported and used, the preparation of teachers, the moderation of their judgements, the way in which evidence from assessment by teachers is combined with that from any external tests, the role that the performance of students has in the accountability of teachers and schools, and how the performance of students is monitored at local and national levels. Unless changes are made in the assessment system, as well as in the methods used, there is a danger that increasing use of teachers' summative assessment could lead to the replacement of high stakes testing by high stakes teachers' assessment – and an increased workload for teachers and an instrumental form of coaching to the assessment criteria.

The simple statement that summative assessment is carried out in order to give information about what students have achieved at particular times belies the different requirements for each of its wide range of uses. These uses range from providing a routine record of all students' progress for schools' records and for reporting on

---

<sup>2</sup> ARG commissioned the review by Paul Black and Dylan Wiliam which led to 'Inside the Black Box' (1998). This was followed by the ARG pamphlet 'Beyond the Black Box' and the leaflet 'Assessment for Learning: 10 Principles (ARG, 2002a).

<sup>3</sup> 'Testing, Motivation and Learning' (ARG, 2002b)

individual students to parents, to the selection of entrants to higher education through a highly competitive procedure. While the concept of summative assessment may be simple, the uses of data from summative assessment are varied and the requirements of different uses make varying demands in relation to reliability and validity of the assessment. Meeting these demands places limits on teachers' autonomy in conducting summative assessment. By contrast, in the case of formative assessment, while the concept is complex, there is a single use in the minds of teachers and students – to encourage deep engagement in learning – and it is under the complete control of the teacher.

## **2.2 The project goals**

The ASF project was set up to address the issues, concerns and reservations that surround teachers' assessment for summative uses.

The overall goals of the project are:

- to clarify thinking by educational professionals, by politicians and by various users of education, about the nature, practice, potential and challenges of assessment by teachers, and
- to provide reports, including recommendations for policy and practice, on the role that assessment by teachers can play in assessment systems.

## **2.3 Procedures**

To achieve these aims the project is attempting to bring together information from the analysis of current practice, the views of participants in, and users of, assessment and what is known from research about the benefits and challenges of assessment by teachers for summative uses. The method chosen was to hold a series of five seminars with invited experts and two consultation conferences with potential 'users' of assessment data.

The project has a Core Group of assessment specialists, practitioners and representatives of policy-making bodies, which provides advice on the programme and participants for the seminars and conferences. It also ensures that information about on-going developments in assessment is taken into account as the project works to achieve its goals. The members of this group (including all members of ARG) attend all the events, with some additional participants invited in relation to the special focus of each seminar.

# **3. Outcomes of the first year's work**

## **3.1 Clarification of terms**

During the first seminar<sup>4</sup> it became evident that a good deal of clarification of terms was needed. Assessment by teachers was interpreted variously, in some cases narrowly, and in others very broadly to include, for example, teachers' participation in setting and marking examination papers. There were also different purposes of assessment by teachers to be considered: some 'internal' to the school (such as for informing students' next teachers or other institutions to which they transfer or their parents) and some

---

<sup>4</sup> Reports of all seminars can be found on the ARG website. See note 1

‘external’ (such as for accreditation or certification of learning, selection, or monitoring school or national standards). Discussion of the definition of teachers’ summative assessment was included in the second seminar programme. As a result the definition in section 4 was proposed.

### **3.2 Developing principles for summative assessment by teachers**

The aims of the ASF project include making recommendations for policy and practice, so it was necessary to devise a framework for evaluating existing procedures for teachers’ summative assessment and for proposing new ones. Decisions of both these kinds, however, require not only that the options be set out, but that the principles guiding the decisions are also made explicit. The project has, therefore, attempted to identify principles for developing teachers’ summative assessment.

In arriving at the principles, which are set out in Section 5, the inputs and discussion in Seminar 2 were found particularly helpful. Seminar 2 brought together information about the ways in which assessment by teachers is featuring in current practice in, and future plans for, assessment systems in the four countries of the UK. In Seminar 3, information about practice outside the UK was brought to bear on the issues surrounding the use of teachers’ judgements in summative assessment. The experience in Queensland, over 32 years, of using assessment by teachers as the basis for the Senior Certificate was described by Graham Maxwell. From the US, the Berkeley Evaluation and Assessment Research (BEAR) system was presented by Mark Wilson. Respondents and other participants noted the cultural and contextual differences that prevent direct application in the UK of assessment procedures from other countries. At the same time some universal factors emerged from the discussion, relating, for example, to the interaction between pedagogy, curriculum and assessment, to the central and crucial role of moderation and to the need for clarity about the relationship between formative and summative assessment. This led to the consideration of issues surrounding how change in systems can be brought about, particularly the political and social changes that are needed in moving towards assessment systems that make more use of teachers’ assessment for summative purposes. Key points from seminars 2 and 3 are summarised in Appendix 1

### **3.3 Describing and analysing methods of summative assessment by teachers**

Summative assessment by teachers can be implemented in many ways. These ways range, for instance, from summarising on-going records of assessments that may also have been used at the time to help learning, to a specified piece of work set according to external guidelines, such as a science practical or a foreign language oral examination, marked by teachers according to a strict scheme. A framework for describing this range of practices was created by first identifying key components of assessment methods and then listing the various ways in which they can be carried out.

The outcomes of the research reviews that have been carried out by the ALRSG (Assessment and Learning Research Synthesis Group, in which all ARG members participate) were helpful in making these decisions. Some relevant points from these

reviews are summarised in Appendix 2<sup>5</sup>. As a result of discussion in ASF project Seminar 2 the components presented in Section 6 were identified.

---

<sup>5</sup> For a more detailed summary of the findings of the reviews and their implications see the forthcoming ARG booklet: Summative Assessment by Teachers: Issues and Opportunities

## Part 2

### 4 Definition of summative assessment by teachers

All assessment, for whatever purpose (formative, summative or evaluative) is a process of reasoning from evidence. The evidence that we use, from students' actions, artefacts and communications, is only an indication of what students know, understand, can do, feel or strive for. We can never be sure that the evidence is a true reflection of cognitive abilities or affective or conative states. So an assessment will always be an estimate, not a precise statement about students' attainments.

The definition of assessment by teachers for summative uses that has emerged from discussions and which is used in the work of the project is:

The process by which teachers gather evidence in a planned and systematic way in order to draw inferences about their students' learning, based on their professional judgement, and to report at a particular time on their students' achievements.

Some elaboration of this may be helpful.

*'The process by which teachers gather evidence in a planned and systematic way'*  
Summative assessment evidence can be gathered in very many different ways. Wilmot<sup>6</sup> has developed an extensive taxonomy of the forms that these ways can take. In section 6 we use a simpler framework for describing assessment methods which depend on teachers' professional judgement. At this point what is relevant is to note that the methods include, for example: data gathered continuously over time, or intermittently, or from specified assignments; data derived from regular work or from tasks that take the form of extended projects, or fieldwork, or presentations, etc; evidence in the form of a portfolio of work, a project report, etc; and moderation procedures that may be internal or external to the school.

... *'to draw inferences'*

Assessment involves making a statement that goes beyond the raw evidence to give this some meaning in terms of a measure of learning. Evidence that is gathered in the process of assessment is compared with some standard or expected achievement. When the standard is the general level of performance expected of students at a particular age or stage, then the assessment is norm-referenced; when it is in terms of the extent to which a particular performance is evident, then it is criterion-referenced; when it is in terms of what the students could do at an earlier time, then it is ipsative or student-referenced. The inference drawn about the achievement of the student will reflect the frame of reference that has been chosen.

... *'about their students' learning'*

The emphasis here is on 'their', thus restricting the meaning of TA in this definition to situations where teachers assess their own students and excluding marking of other

---

<sup>6</sup> Wilmot, J. (2004) Experiences of Summative Teacher Assessment in the UK. A review conducted for the Qualifications and Curriculum Authority.

students' work (apart, that is, from what is involved in moderating teachers' judgements).

... *'based on their professional judgement'*

Judgement is required in making decisions about how evidence relates to the standards or expectations being applied. In some cases, the judgements are holistic, taking evidence across a period of time or an extended piece of work and comparing with assessment criteria. In other cases 'best fit' judgements will be needed perhaps in relation to more than one scale. Professional judgement<sup>7</sup> indicates that understanding of the operational meaning of the criteria and knowledge of the evidence that can be taken as indicators of specific criteria is brought to bear in deciding the extent of match between evidence and criteria.

... *'report at a particular time on their students' achievements.'*

By its nature summative assessment is a summary of what can be said about students' learning at a particular time. This summary can be arrived at through bringing together what has been assessed at intervals during the time covered by the report, or can be based on a snap-shot at the reporting time, or a combination of these.

## **5 Principles of summative assessment by teachers**

We believe that assessment by teachers can and should have a significant part in the summative assessment of student achievement. We are not claiming that assessment by teachers is of itself sufficient for all the purposes that summative assessment has to serve. For several summative uses, teachers' assessment needs to be complemented by information from externally produced tasks or from sources outside the school that provide evidence of achievement of non-academic goals. Our focus here, however, is on the part that teachers' assessment can take in wider reporting system.

### **5.1 Overall goals**

The overall goals in developing procedures and systems for using summative assessment by teachers are:

1. that summative assessment is seen by teachers and students as a positive contribution to teaching and learning
2. that teachers' assessment is used to provide a dependable contribution to summative assessment
3. that teachers' assessment enables achievement across the full range of goals of learning to be reported
4. that the components of any method of summative assessment by teachers should be coherent, internally consistent and suited to its particular use.

### **5.2 Principles**

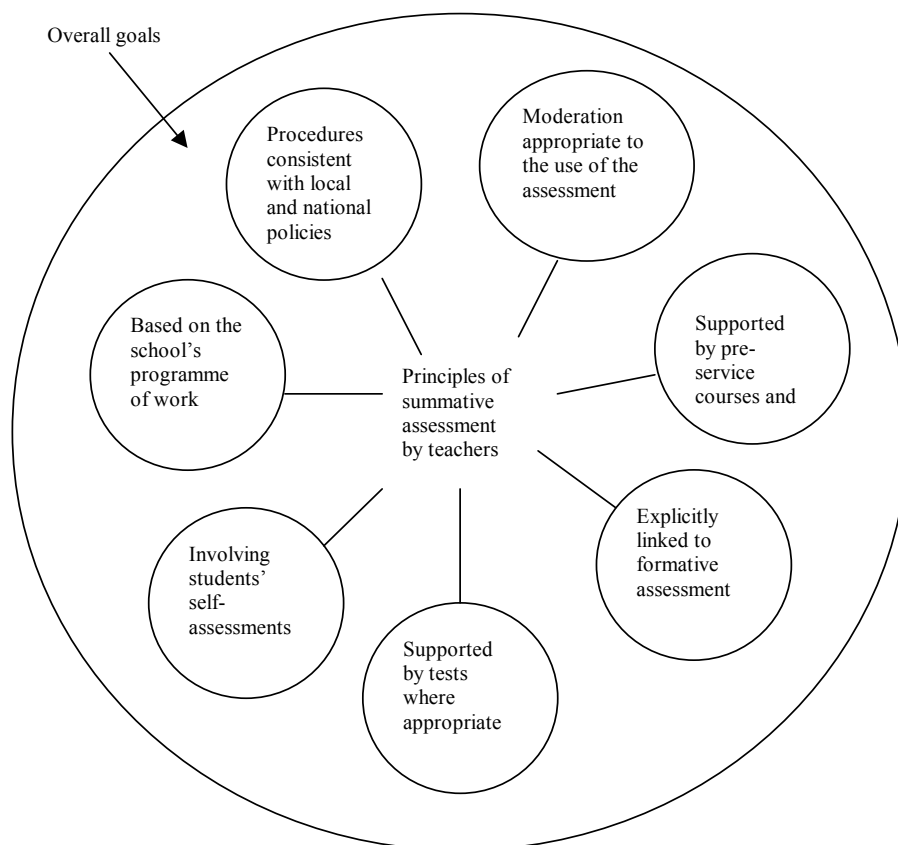
To achieve these goals the following points from research practice are proposed as principles for guiding the development of procedures and systems. Their arrangement

---

<sup>7</sup> According to Hargreaves the exercise of professional judgement means the capacity to make discretionary judgements in the interests of improved student outcomes and transformations (Hargreaves, 1994). The new professionalism: the synthesis of professional and institutional development. *Teaching and Teacher Education*, 10 (4) 423 - 438

in Figure 1 is intended to indicate that there is no order of priority among them. There follows a rationale and discussion of the meaning of each one.

Figure 1



### ***5.2.1 Teachers' summative assessment should involve moderation designed for quality enhancement appropriate to the purpose and uses of the assessment***

The term 'moderation' is used here to refer to the part of the assessment system that is designed to validate the data from teachers' judgements. This doesn't mean that it is a procedure that only affects the outcome, serving a quality control function. Rather we see moderation as having an impact on the process of assessment by teachers as well as the product; that is, having a quality assurance function. The dependability of assessment originates in what happens when teachers plan and conduct assessment; it is not just a matter of adjusting the product. Thus moderation has to be conducted in a way that affects the processes of planning and implementing assessment, which in turn depend upon teachers' understanding of learning goals and of the criteria indicating progress towards them. Moderation should be planned to guard against the errors and bias that research shows exist in teachers' un-moderated judgements<sup>8</sup> by: prevention, as far as possible, at the point when judgements are made; by detection and

<sup>8</sup> See the ALRSG review of research: A Systematic Review of the Evidence of the Reliability and Validity of Assessment by Teachers for Summative Purposes. Available from the web at: [http://epi.ioe.ac.uk/EPPIWeb/home.aspx?page=/reel/review\\_groups/assessment/review\\_three.htm](http://epi.ioe.ac.uk/EPPIWeb/home.aspx?page=/reel/review_groups/assessment/review_three.htm)

correction of the resulting judgements; and by enhancing the quality of all the interactions and procedures involved

The details of how quality enhancing moderation is carried out vary with the uses of the assessment data. These fall into two groups: those for 'internal' school use and those for use by agencies, individuals or bodies external to the school. Assessment for internal school uses means assessment that is conducted by teachers to keep a record of students' achievement and to report on the individual achievement to the students' next teachers, their parents and the students themselves. Such assessment may take the form of a summary of achievement over time based on regular work, occasional tests, end of year school examinations, etc. Moderation procedures are devised and controlled by the school, conforming to any external guidelines or requirements as may apply. The procedures may include: school, year-group or subject department meetings to discuss criteria for awarding grades or levels; development and use of a set of school examples of the application of criteria; discussion of specific student profiles or checks on marking.

External school uses are those where teachers' assessment is used in part or whole as evidence for certification, selection or deciding levels of achievement against criteria that have been determined externally. The information is reported to examination boards, local authorities, employers or other educational institutions and may also be reported to parents and students. For these purposes, there is a greater need than for internal school uses for the rigour of the moderation procedures to be evident and explicit. This may entail a formal accreditation of the school's assessment and moderation procedures or accreditation of individual teachers by an external body, or the direct involvement of the external body in the process of moderation.

### ***5.2.2 Teachers' summative assessment should be supported by pre-service courses and continuing professional development***

Teachers' assessment takes advantage of the detailed knowledge of students that teachers build up through frequent contact over a period of time. Teachers' confidence in using this knowledge for summative assessment depends on a firm understanding of the purposes and processes of assessment and recognition of possible sources of bias and error. There is research evidence that information that is not relevant to a particular kind of achievement may influence their judgements (see Appendix 4 and Working Paper 2). For example, teachers' judgements of young children's achievement may be biased by the children's behaviour in class. Such bias can be minimised by training designed to address these problems.

Courses should provide more than information about what is involved in collecting and judging information for assessment. They should give trainees and teachers experience of generating as well as using criteria so that links are clearly made with learning goals. Such experience leads to greater insight into the meaning of both learning goals and the expected standards or quality of work. There should be opportunities to discuss judgments made across a range of student work in order to understand the difference between assessing individual pieces of work and summarising achievement over a period of time. The relationship between formative and summative assessment should also be considered so that, when teachers are using information collected over time they can use it to help learning as well as to assess achievement, as suggested in 5.2.3. Courses with these features improve the quality of teaching and learning as well as the dependability of teachers' summative assessment.

### ***5.2.3 Teachers' summative assessment should be explicitly linked to formative assessment***

Teachers' summative assessment conducted by collecting information over a period of time has two considerable advantages. First, it avoids the single, end of course test or examination, which causes anxiety in students, particularly girls and lower achieving students, leads to preparation and practising for tests can distort teaching and learning (see ARG, 2002)<sup>9</sup>, and pins everything on performance at a particular time which may or may not be typical. Second, it enables the evidence of achievement at earlier stages in the course to be used formatively to assist further progress in the later stages. The consequence is that more weight should be given to the performance at the later than at the earlier stages. However it is important that the different purposes of formative and summative assessment are recognised in the way they are carried out. Although the same information is used for both purposes, the purposes are different and the way in which judgements are made must reflect this difference. Making the relationship explicit will help to avoid bias in summative assessment and protect the integrity of formative assessment. At the same time it will recognise the inter-relationship between assessment the two purposes.

Information gathered about the progress of a student towards the goal of a particular piece of work (which may be chosen to suit the individual needs of the student) is used formatively when it is used to inform feedback about the next steps the student needs to take. The feedback should indicate these next steps but should not, if it is to be really useful to the students, be in terms of a judgement about what level or grade has been achieved. For summative assessment, the information from work over a period of time (including ephemeral evidence) is brought together and judged in relation to grade or level criteria that are the same for all students. Teachers should be able to relate student-specific goals to the goals for all students. But they should not be constantly judging students against the latter because this may well be discouraging, particularly for lower achieving students and would turn formative assessment into a succession of summative judgements. Summative assessment should be conducted only as and when it is needed for a clear use that is evident to the students and the intended users.

### ***5.2.4 Teachers' summative assessment should be supported by externally devised tasks or tests where appropriate***

In some circumstances and for some learning goals regular work may not provide all the evidence needed for summative assessment. Inexperienced teachers, particularly, may need help in identifying how goals such as problem solving, critical reasoning and the application of knowledge in new situations can be identified. Well designed examples of tasks, with marking schemes, designed to assess these learning processes and outcomes can help teachers to ensure that they are identified and included in the evidence used by teachers in their assessment. Their use would be optional and, when used, would provide pupils with interesting and relevant experiences, exemplifying for teachers how opportunities for relevant learning can be provided.

We emphasise that such tasks are not to be seen as separate measures to be set beside teachers' judgements. Nor are they intended to confirm teachers' judgements. Rather they are part of the evidence that teachers can use, if they choose to do so, to ensure that all intended goals are taken into account in their assessment.

---

<sup>9</sup> ARG (2002) Testing, Motivation and Learning.

### ***5.2.5 Teachers' summative assessment should involve students' self-assessment of academic and non-academic achievement***

Students' self-assessment has a key role in formative assessment because students are in the best position to know what they really understand and what skills they feel confident in having developed. For the same reason, students have a role in summative assessment. The more students can be involved (in ways appropriate to their age and maturity) the better they will understand the meaning and purposes of the summative assessment and see it as a necessary part of their education. For this to happen teachers need to share with students the learning goals and the criteria by which achievement is judged. As with formative assessment, this sharing can help students' engagement in their learning. For summative assessment, however, the sharing of learning goals and criteria has to be done in the context of reaching a judgement progressively over time, as noted in 5.2.3. Students should be using self-assessment formatively, rather than participating in a series of short-term summative judgements.

As stated at the start of this section, assessment by teachers is part of a system for reporting student progress that includes both academic and non-academic goals. Thus the achievement of students that need to be recorded and reported for some uses include progress towards goals that are achievable through activities outside the classroom and school as well as within. These goals include, for instance, care for the environment, consideration for others, teamwork. Evidence of these achievements from outside the classroom will need to be gathered by the students themselves, using the skills of self-assessment developed in the school context.

### ***5.2.6 Teachers' summative assessment should be based on the school's plan for achieving agreed curriculum goals***

Using teachers' assessment for summative purposes has to be seen in the context of there being substantial responsibility at the school level for the teaching, learning and assessment of students. There are two particular reasons for this. First, it is reasonable to expect teachers to have some part in deciding what is taught, and how, if they have the responsibility for assessing it. They are in the best position to decide how to pursue particular goals for their students, using the local environment and building on particular strengths of the school's amenities. Second, research shows that teachers' assessment is more reliable and valid when based on a programme of work determined by the school. The more involved teachers are in decisions about the programme of work the better grasp they will have of, and commitment to, its goals. In turn this will mean that assessment is more reliable. However, these goals have to be educationally worthwhile and consistent with nationally and locally agreed curriculum guidelines of requirements.

### ***5.2.7 Teachers' summative assessment should use procedures consistent with local and national assessment policies***

A similar point to the one made about top-down approaches to the programme of work can be made in relation to the procedures used for summative assessment by teachers and its moderation. Externally imposed procedures need to be communicated, interpreted and understood if they are to be implemented effectively but it is also clear from research that teachers are more readily able to operate procedures they have helped to decide. It is, of course, essential for the moderation procedures to be consistent with local and national policies so that the same judgements have the same meaning across schools. However, this does not mean that the procedures for collecting information have to be identical, any more than teaching approaches have to be

identical in order for students to achieve the same goals. The essential feature of dependable teachers' summative assessment is that judgements are made against the agreed criteria and that moderation ensures that an adequate sample of evidence is reliably judged against these criteria.

The methods of gathering information about student achievement vary widely, from teacher-made tests to review of regular work, to computer-based assessment. Guidance should be given to teachers in relation to the appropriateness of each of these according to the goals of the work and the use to be made of the data.

### **5.3 Managing assessment by teachers in a summative assessment system**

#### **5.3.1 Teachers' tasks**

At the centre of the system is *the individual teacher* who is making judgements about his/her students' work. The teacher assessor is managing students' learning activities and making judgements about their attainments as part of his/her professional responsibilities. S/he is likely to be making judgements about a diverse range of work that was not completed under the same type of tightly defined conditions as apply to externally set tests or examinations. Moreover, teachers have access to other information about the students, such as their general behaviour and attitudes, which can introduce 'noise' into the system. Hence, if the teachers working within the same assessment system in a single institution are to ensure that their judgements are consistent with each other there will need to be internal moderation procedures, for which a designated person, an internal moderator or moderating group may be made responsible for a subject of group of subjects.

Any system, whether based on teachers' judgements or not, needs to ensure that it effectively communicates expectations. There are some parallels in this respect between the management of externally marked tests and schemes involving teachers' assessments. Unambiguous definitions of the type of task or of the assessment criteria can contribute to consistency of practice across the scheme. There is, however, an abundance of evidence that, for effective use of the discretion available to teachers, those involved need to meet with each other. In teachers' assessment there is often no common task and specific set of expectations of students such as is usually the case in the marking of examination scripts. Instead each teacher is faced with making judgements, in relation to published criteria, of a typically wide range of student responses to a more loosely defined set of tasks. In those circumstances, individual teachers understandably need opportunities to discuss, both with each other and with those in moderation roles, how to interpret and to apply the criteria when judging their students' work. The nature and extent of such opportunities for discussion is a crucial consideration in the design of a system for using assessment by teachers for summative purposes.

#### **5.3.2 Quality management**

If a system based on teachers' assessments and extending beyond a single school is to achieve acceptable levels of consistency in the ways that judgements are made within it then it needs effective quality management procedures. These will include moderation of the procedures that are in place to prevent error and bias and moderation of the outcomes of assessment to detect any inconsistency in the way that criteria have been

applied (see 5.2.1). The way in which moderation for these purposes is conducted should be decided with a view to quality enhancement. In brief the elements are:

- *Quality control: monitoring of outcomes.* There is a range of options available<sup>10</sup> for the monitoring of outcomes from schemes. The relative merits of those options can be debated but it is likely that any such scheme will wish to make use of checks on patterns of student performance with a view to identifying and investigating any apparent anomalies.
- *Quality assurance: monitoring of procedures.* The definitions of roles and responsibilities, together with specific arrangements for communication within a system, will give rise to a set of operating procedures for a system based on teachers' assessments. Just as those responsible for examination-based schemes typically put in place checks on whether procedures are being followed, so those responsible for teacher assessment based schemes need to have checks in place. Such checks are as important in establishing the confidence of those working within the scheme in its robustness as they are in building wider public confidence in the credibility of systems where teachers are making the crucial initial judgements.
- *Quality enhancement.* One of the most striking features of the more sophisticated schemes<sup>11</sup> that are based on teachers' assessments is their continuing commitment to quality enhancement. Both courses that enable teachers or trainees to come together to develop and discuss assessment criteria and those where assessments are moderated through discussion are powerful vehicles for professional development. They allow teachers to pool and develop their expertise, not only in relation to the assessments they are making but also in relation to the ways in which they are implementing the curriculum and providing contexts for students' learning. Quality enhancement considerations should be taken into account from the outset when designing a system so that, for example, any arrangements for face-to-face communication are understood to be at least as much about quality enhancement as about quality assurance and control.

## **6. Model for describing and evaluating methods of summative assessment by teachers**

### **6.1 Components of summative assessment methods using teachers' judgements**

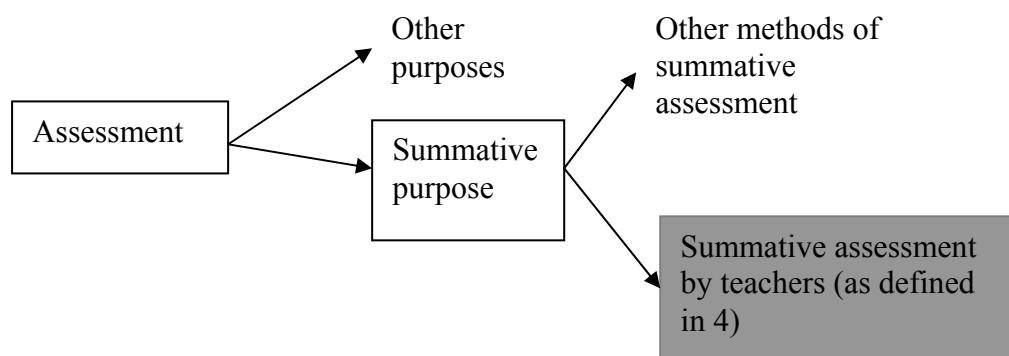
Summative assessment refers to assessment for one of the main purposes of assessment; the others being variously described as formative, evaluative and monitoring. In meeting the overall purpose of summative assessment, information can be gathered and judged in a range of ways, one of which is described here as assessment by teachers (according to the definition in section 4). Again, there are other methods for conducting these processes, such as by teachers administering tests or tasks that are externally marked, external tests and examinations, etc. Figure 2 puts the focus of our concerns here in the wider context of other purposes and other methods of assessment.

---

<sup>10</sup> Harlen, W (Ed) (1994) *Enhancing Quality in Assessment*. London: Paul Chapman Publishing

<sup>11</sup> For example the Queensland Senior Certificate. See papers presented at the 3<sup>rd</sup> seminar at [www.assessment-reform-group.org/ASF.htm#reports](http://www.assessment-reform-group.org/ASF.htm#reports)

Figure 2



In order to examine the components of summative assessment by teachers and their impact on the dependability and acceptability of various approaches, we have developed an analytical model. The aim in doing this is not just to describe approaches, nor it is to provide an exhaustive list of all possibilities. Rather it is to develop a framework, or model, for critical analysis of different existing approaches and for developing others. Such a model should be as simple as possible, whilst including all the factors that are known to influence the effectiveness and reliability of the assessment.

A dependable method of summative assessment within which teachers' judgements are central to the outcome has certain essential components. From an analysis of such methods it is suggested that these are:

- Uses (relating to whether this is for internal school uses, such as records handed from teacher to teacher, for parents or for external use, such as certification, accountability)
- Type of task (relating to the situations in which teachers assess students as required by the school or other authority or decided by the teacher)
- Nature of criteria applied (relating to the detail and structure of criteria or marking schemes, since all the approaches are criterion-related)
- Form of judgement made by teacher (relating to the form in which the judgement is expressed and recorded)
- Guidelines and/or training (relating to the training in assessment for teachers given or recommended nationally or locally)
- Moderation (relating to the procedures set up to provide confidence of users in the assessment outcomes).

Various ways in which these components can be put into practice are set out in the model in Figure 3. The variables are not, in most cases, mutually exclusive. Various current procedures for teachers' summative assessment can be described by combination of these variables. By bringing research findings and evidence from successful practice to bear on assessment for a particular use, the profile of a dependable method of summative assessment by teachers may be identified. For instance, in the case of the 'internal' use for reporting to students and their parents the profile might be:

Type of task:

- Regular work or embedded tasks created by teachers

Criteria applied

- Detailed but not task-specific criteria (multiple scales), linked to development towards learning goals

Type of judgement made by teacher

- Levels of achievement against specified criteria (profile)

Guidelines/training

- Training procedures created and coordinated locally

QA/Moderation

- Specified within-school moderation procedures

For external purposes, there is a wider range of options, for example

Type of task

- Regular work or embedded tasks created by teachers
- Embedded tasks selected from range externally provided
- Specified embedded tasks
- Tests/tasks created by teachers
- Tests/tasks selected from external pool or bank
- Tests/tasks prescribed externally

Criteria applied

- Detailed non-task-specific criteria (multiple scales), linked to development towards learning goals
- Detailed task-specific criteria (multiple scales), linked to development towards learning goals

Type of judgement made by teachers

- Mark or score based on criteria or marking scheme
- Levels of achievement against specified criteria (profile)
- Overall level or grade achieved

Guidelines/training

- Training procedures created and coordinated nationally

Moderation

- Specified within school moderation procedures
- Specified inter-school moderation procedures
- Inspection of procedures/evidence by external moderators
- Teacher/school/centre accreditation

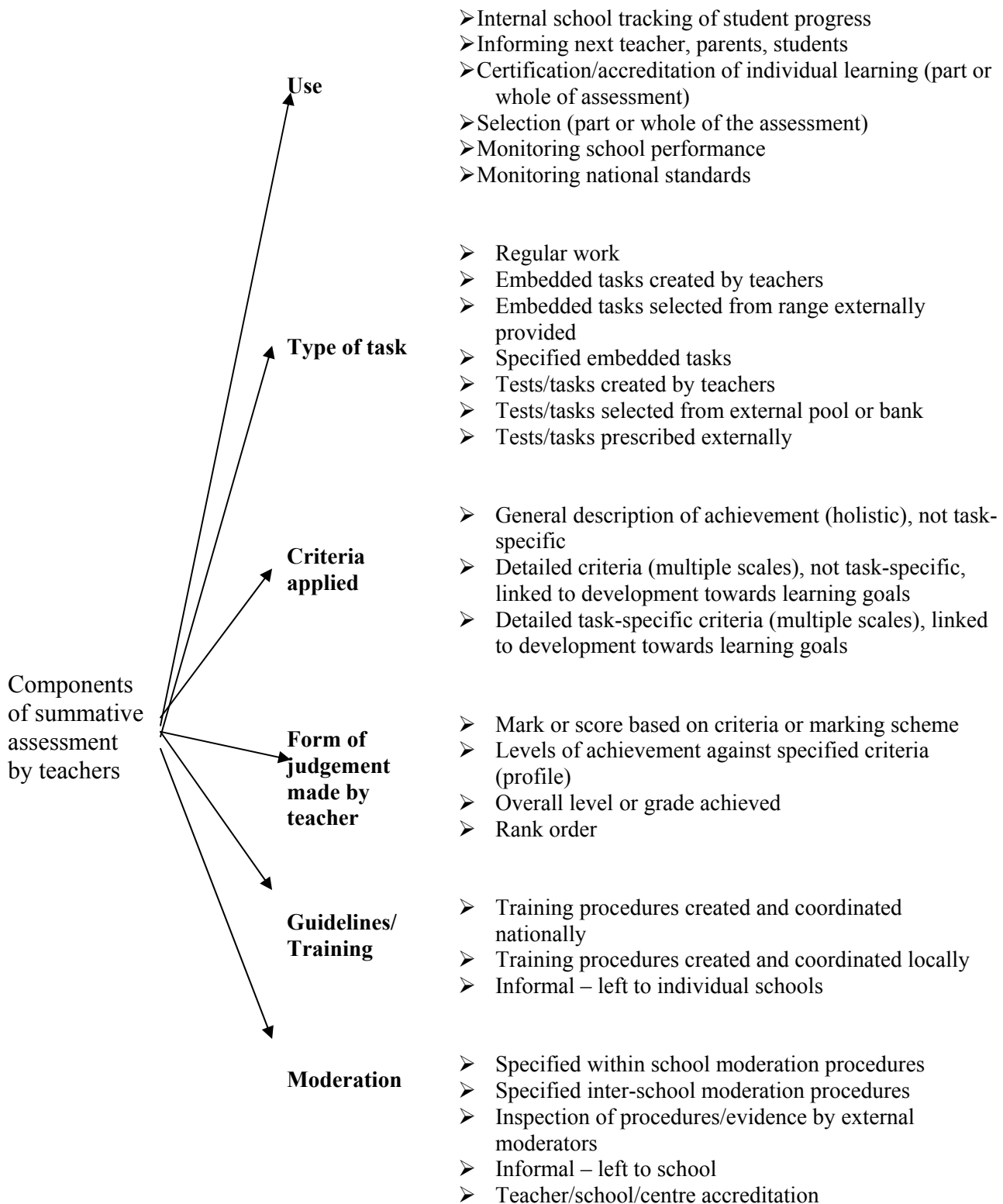
## 6.2 Representing the dynamic relationships among the components

Figure 3 represents the variety of components that are relevant to analysing approaches to teachers' summative assessment. However, in order to understand how a system of assessment operates it is necessary to consider the dynamic interaction among the teachers, the assessment objectives and the procedures used in assessment. For example, in a classroom in which a teacher has developed the practice of strong formative assessment, which feeds in to summative assessment at the end of a topic, module or specific period of time, certain classroom procedures will predominate. These are likely to include honest interactions about their learning both among students and between students and teacher, so forming a distinctive and lively community. The teacher will also treat students' written work in such a way that the feedback is available for parents and colleagues for inspection. If the teacher faced a new requirement, to produce prescribed pieces of students' coursework to contribute, via moderation procedures, to summative assessment, the classroom dynamics would alter.

At certain times the summative assessment would require specific and constrained procedures to ensure fairness among different students, classes and schools. At these times the open interactions with and among the students might have to be suppressed, so altering the community and conflicting with the expectations set up by the formative assessment procedures.

The overall point here is that to understand, and so perhaps to design and study how any new ways to enhance the quality of teachers' summative assessments might work out in practice, a theoretical framework is needed to deal with the dynamics of interaction and change between the various groups in the contexts of their assessment activities.

Figure 3 Variables within the components of summative assessment methods using TA



## **Part 3**

### **7. Reflection, synthesis and implications**

#### **7.1 The evidence and policy context**

Education policy should be informed by evidence. However, for many of the aspects of assessment discussed in the project seminars the evidence does not give clear guidance. For instance, while formative assessment has the potential to raise standards, implementation can be so superficial that potential benefits are not realised. Summative assessments by teachers can satisfy requirements for fairness and rigour, but there are many examples of innovations where they have failed to do this.

Experimental studies can establish that improvements are attainable. However, that is a necessary but far from sufficient condition. Large-scale implementation is fraught with problems, and there are some who believe that ‘roll-out’ and ‘scaling-up’ are tricky metaphors that give a veneer of simplicity to approaches that treat teachers as technicians and institutions as homogeneous.

Even if such problems could be tackled successfully, and all the necessary evidence and the strategies for authentic implementation of innovations were in place, optimum policy for assessment issues would still be hard to achieve. It may be necessary to convince politicians, the media and the wider public that their ‘common sense’ view of what is desirable is no more reliable a guide in education than it is in (say) medicine.

The complexity of assessment and its related impacts attracts public scepticism, even when well-researched. Assessment serves several different purposes in education and policies can privilege one of these at the expense of others. To prefer any one gain at the expense of a corresponding loss is to make a decision about values – and there is no escaping the fact that assessment policies are so close to the heart of educational action that the most important decisions will involve choices determined by personal and social values.

Furthermore, whilst an analytic approach, in which learning, pedagogy, certification, management, accountability and so on are all each analysed in separate detail, is inevitably necessary, it is bound to be far from adequate. All such elements interact, closely and in complex ways. Thus it is not surprising that policies may fail to achieve their purposes and can give rise to unintended and unacceptable effects on practice. A systemic approach is essential – but it must be an approach that can deal with the dynamics of rapid social change.

Aware that there are no easy answers, the following discussion leads us to recommendations relating to:

- The learning to be assessed
- Different sources of information for summative assessment
- The use of teachers’ time
- The relationship between summative assessment and formative assessment
- Professional development for dependable summative assessment

- Summative assessment of students and school accountability
- The interaction of assessment, curriculum and pedagogy

## 7.2 Implications

### 7.2.1 The learning to be assessed

A central argument for change in current systems of summative assessment arises because they do not adequately reflect the range of educational outcomes that are constantly affirmed as needed in a world of rapid social and technological change. These outcomes include higher order thinking skills, the ability to adapt to changing requirements and communication modes, ‘fluid’ learning that is applicable in a range of contexts, the ability to work in groups, in addition to curricular knowledge and understanding. This happens because the use of individual student data for accountability and monitoring requires priority to be given to reliability and so to forms of assessment, ie tests, that are assumed to yield ‘accurate’ data, despite evidence to the contrary<sup>12</sup>.

Further the high stakes attached to the results encourages transmission modes of teaching that promote shallow and superficial learning rather than deep conceptual and procedural understanding. Research also shows that high stakes tests (and teachers’ own tests that emulate these) have a negative impact on some students’ motivation for learning and their engagement with learning in other than an instrumental way to achieve summative targets. There is also firm evidence that dependence on tests inhibits the development of practices of formative assessment that are proven to raise the levels of achievement and reduce the gap between the higher and lower achieving students. In sum, not only do many aspects of summative assessment systems that are based on tests have detrimental impact on students and teachers but they fail to provide valid information about students’ achievements.

Summative assessment is a necessary part of education and serves important purposes, but systems with these failings need to be reviewed and modified. While assessment by teachers is not free from all of these failings, it suffers from fewer of them. It has the considerable further advantage of promoting teacher professionalism and continuing quality enhancement.

#### *Implication 1*

A comprehensive summative assessment system must be capable of providing information, based on dependable judgments, about how much and how well students have developed a wide range of cognitive competencies and affective outcomes.

### 7.2.2 Sources of information for summative assessment

From the students’ perspectives, their school work is but a part of their learning experience. Outside school they continue to develop as people and as learners. For a full record of their progress, therefore, information needs to be gathered from activities both within and outside the school.

---

<sup>12</sup> See, for example, Chapter 7 in J. Gardner (Ed) (2006) *Assessment and Learning*. London: Sage

In relation to the information from within the school, for reasons summarised above, we propose that this should be based as far as possible on teachers' assessment. In some circumstances it will be necessary for teachers' judgements to be supplemented by information from externally devised tasks or tests. The reason may be that these tasks present the best way of providing evidence of more complex cognitive skills or learning with understanding. A well designed set of assessment tasks available for teachers to use have several benefits. They exemplify for teachers the situations in which skills and understanding are used and thus guide them in developing their own embedded assessment tasks. They are also of particular benefit to newly qualified teachers and those who need to build their confidence in their ability to assess students.

For information from outside the classroom, students themselves are in the best position to report this. Out of school experiences such as involvement in family, community or environmental care, sport, music, dance, fund-raising, etc can all provide evidence of development of responsibility, creativity, team-working skills, self-discipline, and so on that should be part of an on-going record of achievement. Involving them self-assessment in the classroom develops skills that they can use in reporting their achievements in relation to non-academic goals.

These different sources of information – teachers' assessment, performance on external tasks or tests and reports of non-academic achievement need to be seen as complementary to each other. There should be no attempt to combine them, for to do so would require the relative value of one against others to be judged and quantified.

### ***Implication 2***

Summative assessment systems should be designed so that information appropriate for specific purposes (i.e. fit for purpose) can be collected and reported.
-------------------------------------------------------------------------------------------------------------------------------------------------------------

### ***Implication 3***

Teacher should have access to optional tasks assessing skills and understanding, which they can use to assist them in making judgments across the full range of learning goals.
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### **7.2.3 The use of teachers' time**

Our initial argument that teachers are in the best position to provide valid data about the full range of goals of learning has been strengthened by presentations at the project seminars and consultation conferences. Teachers are all the time exercising their professional judgement about how to progress students' learning and routinely conduct summative assessment for 'internal' school uses (see 5.2.1). They can extend this to serve more formal 'external' summative assessment uses, given the necessary structure, support, training and time. Indeed those who are also involved in the assessment of individual students for external certification clearly have confidence in their own judgements. This is shown in the large number of appeals against examination grades and the high proportion of judgments in favour of teachers' grades in these appeals.

Teachers can gather information about learning processes as well as learning outcomes. The range of methods that they can use makes 'teaching to the test' a meaningless concept and supports a variety of teaching modes. They can take advantage of practising formative assessment by using the information gained to summarise learning at a particular time, as well as to help learning at all times.

However turning the information that teachers can gather into reliable data, as required for summative assessment, is not without many problems. Attention has to be given to procedures and training to meet the requirements for dependability. That this involves teachers' time, including a good deal of non-teaching time, is not being ignored. Some quite fundamental redeployment of teacher resources is likely to be required. The benefits of moderation procedures are, in theory, considerable although perhaps only realised through participation. As noted earlier, and supported by research, teachers' involvement in discussion of the students work not only clarifies their understanding of learning goals and assessment criteria, but also provides insight into students' difficulties and give ideas for teaching to overcome them. Teachers will also come to recognise the time saved when reliance on external tests and examinations is decreased. Crucially, it is essential to ensure that teachers recognise the benefits of making greater use of their judgments and that their concerns about increase in workload are fully addressed.

#### ***Implication 4***

Summative assessment systems should be designed to enhance the overall validity of assessment and support learning with no additional burden on teachers and students.
------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### **7.2.4 The relationship between summative assessment and formative assessment**

Summative assessment has to be considered, not in isolation from, but in the context of assessment for other purposes. The relationship between formative assessment and teachers' summative assessment is particularly important since both are in the hands of the teacher and have impact on students and teachers and their interactions. The relationship between formative assessment and summative assessment is a complex matter.

The starting point in our thinking is that formative assessment is an important part of teaching and has to be protected. A strong reason for this view is that classroom practice should reflect not only the full range of goals of an up-to-date education but also the pedagogy that is suited to achieving these goals. This pedagogy reflects modern views of learning that emphasise the development of understanding by learners themselves rather than the transmission of knowledge by teachers. Formative assessment has an essential role in learning viewed in this way. A key component of this, which is of relevance to summative assessment, is the role of the student in reviewing their work and reflecting on their learning. For this, students need to know the goals of their learning and the criteria by which it is assessed. When these criteria are in the form of a series of statements forming a hierarchy of levels, they can serve both formative and summative assessment purposes. The explicit sharing of goals and criteria enables students to know where to invest their effort. Moreover, being involved in self-assessment as part of learning means that students are not faced with summative assessment results that are unexpected, avoiding damage to their self-esteem and motivation for learning.

However, there is firm evidence that emphasis on external tests, particularly when high stakes are attached to the results, is associated with teachers making very little use of assessment formatively to help learning. This provides a strong argument for developing summative assessment systems that give assessment by teachers a significant role. Although exchanging teachers' assessment for tests will not

necessarily diminish the impact of high stakes use of assessment data, it does mean that teachers will be gathering information, which they have the opportunity to use formatively. Also, many of the steps that can be taken to reduce the negative impact of high stakes assessment (whether by teachers or by tests) are ones that favour the implementation of formative assessment. These steps include providing non-judgemental feedback that helps students know where they are in relation to learning goals; teachers sharing with students the reasons for and goals of assessment; teachers of using assessment to learn more about their students and reflecting on the adequacy of the learning opportunities being provided; teachers and students placing less emphasis on comparisons among students and more on individual development; and students being helped to take responsibility for their learning and work towards learning goals rather than performance goals.

#### ***Implication 5***

Summative assessment must be in harmony with the procedures of formative assessment, supporting the use of assessment by teachers and students to help learning.
------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### **7.2.5 Professional development for dependable summative assessment**

All summative assessment needs to be valid, in that it captures the achievement it purports to reflect, and as reliable as necessary for the use to which the result is put. This combination of validity and optimum reliability is described as dependability. The procedures needed to ensure dependability concern both the process of assessment and its product, the result.

In relation to the process, the research on the reliability and validity of teachers' summative assessment reveals examples of the bias and error that can occur in the results in certain circumstances. It also indicates actions that can be taken successfully to minimise error in these assessments. Most error occurs when criteria of assessment are either absent or so vague that they are interpreted differently by teachers, or when judgment is influenced by knowledge of student characteristics that are not relevant to the particular achievement being assessed. There is evidence that these errors can be minimised, in the first place, by having detailed criteria, where possible describing progressive levels of achievement. In the second place, professional development and training that is designed to ensure understanding of criteria before they are used and of how to detect and avoid sources of bias.

Teachers' understanding of criteria is helped by involving them, as far as possible, in developing criteria. Moderation in the form of discussion among teachers of specific examples, engages teachers in close scrutiny of the goals of learning and how these are represented in the example of work. It is therefore of great benefit to teaching and learning, and for the teachers' subsequent assessment using the criteria. These points apply both to 'internal' and to 'external' uses of the results.

The importance of professional development has been explicit or implicit in the evidence and experience reviewed by the project. It is clear that specific attention needs to be given to helping teachers to gather relevant evidence and to use criteria reliably. In addition, there are broader aims of professional development. These include recognising the part that teachers can take in minimising the negative impact of summative assessment on students, for instance in the way that tasks are presented, and

understanding the process of summative assessment and how it can be used to review teaching practice and to help learning.

In relation to the products, summative assessment systems should include checking procedures, whose formality will depend on the use of the assessment result. For internal school purposes a check on patterns of student performance is likely to be sufficient to detect systematic bias that can then be addressed in moderation meetings. For external use, a variety of procedures are already practised, for instance in awarding vocational qualifications as well as in relation to coursework components of end of secondary school external examinations. These include the assessment of samples of work being checked by external agencies. However, where teachers' assessment may be based on judgements of work over a considerable time and the context in which it is undertaken has a significant role, accreditation of the school or individual department or teachers may be more appropriate. Where results are important for the students' future, confidence in teachers' judgments may be underpinned by use of externally devised tasks or test, which is not an assessment in its own right but is used to adjust levels without affecting students' ranking.

#### ***Implication 6***

Both pre-service and in-service professional development should specifically address the development of teachers' understanding and skills of assessment for different purposes, potential bias in teachers' assessment and help teachers to minimise the negative impact of assessment on students.

#### ***Implication 7***

Robust and permanent procedures for moderation of teachers' judgments are needed to ensure that teachers' summative assessment can be used in giving a valid and reliable account of student learning.

#### **7.2.6 Summative assessment of students and school accountability**

A recurring theme throughout our study of research and practice has been the impact of the 'stakes' attaching to a summative assessment – whether the assessment is by teachers or an external agency. The issue is clear and well rehearsed: if there is a high 'price', either for teachers or for students, for achieving a certain level or grade, there is tendency for effort to be focused too narrowly on what is needed to achieve that level or grade. In the case of external tests, this leads to 'teaching to the test' with all that follows from that for the curriculum and for teaching methods. In the case of teachers' assessment it may mean 'giving students the benefit of the doubt' and judgements being more lenient than they should be, or relaxing rules so that students have more help than is intended. Importantly, this leads to distrust of teachers' judgments in favour of what are assumed to be more 'objective' assessment methods: externally marked tests.

Therefore the mode of assessment cannot be divorced from the use that is made of the results. When results have high stakes for students, such as for entrance to higher education, it is clearly important for moderation to be rigorous and explicit. It is equally important that these results are not used for other purposes that put pressure on teachers, particularly school evaluation. It is noteworthy that a feature of the successful

use of teachers' assessment in the Queensland Senior Certificate is that the results are not linked to accountability.

### ***Implication 8***

Data from summative assessments should be reported and interpreted in the context of the broad set of indicators of school effectiveness. Every effort should be made to reduce the priority given to high stakes summative assessment, since this leads teachers to reduce their own assessment tasks to routine events and restricts students' opportunities to learn from the assessment of their work.

### **7.2.7 The interaction of assessment, curriculum and pedagogy**

Research evidence shows that teachers can use criteria consistently across a range of different activities related to a particular goal when using criteria that are sufficiently specific to that goal to enable the presence or absence of relevant aspects of performance to be identified. These are goal-specific, not task-specific, criteria and so they allow achievement of a goal to be assessed through a range of classroom activities. As noted above, there are distinct advantages to having criteria that specify performance at a series of levels relating to a particular goal. Use of such developmental criteria not only facilitates a 'best fit' judgement across a range of work, it provides immediate indication of next steps in learning relating to that goal. They need to be integral to the curriculum and specified in a way that reflects consistency among assessment, curriculum and pedagogy.

The research evidence that external summative assessment controls teaching methods and restricts students' learning opportunities is a key factor in making the case for greater use of teachers' assessment for summative uses. But teachers' freedom to adopt a broad range of teaching methods, reflecting modern views of learning, also requires a greater freedom to determine their school's programme of work. Research shows that both teaching and assessment benefit from teachers using their professional expertise in these matters, rather than being controlled by top-down requirements.

### ***Implication 9***

Attention and resources must be given to creating developmental criteria, which indicate a progression in learning related to particular goals which can be applied to a range of relevant learning activities. Procedures to enhance the teachers' own contribution to summative assessment, should reflect this priority.

### ***Implication 10***

Whole school planning and self-evaluation should involve teachers in supplying summative data in order to facilitate their understanding of and commitment to, learning and curricular goals and to enhance the dependability of teachers' assessments

## **Appendix 1 Key points from practice in different countries**

- The political and social context has a central role in determining the extent to which teachers are given a greater role in summative assessment. For instance, the confidence the society has in teachers and the priority given to students being able to apply their learning to authentic problems will influence policy decisions.
- The confidence of users of summative assessment information and of teachers themselves in teachers' judgements has been severely reduced during the past 15 years in the UK and needs to be restored as part of the process of change.
- Teachers have become dependent upon external tests, which are sometimes used as control mechanisms as well as for reporting on students' progress; it will take time to replace these practices by ones where teachers are able to use on-going assessment to replace external tests.
- Top-down approaches to teachers' assessment, whereby teachers are controlled by strict guidelines and quality assurance procedures are not as effective as are approaches involving teachers in building up necessary procedures using their professional expertise.
- Effective moderation and professional development are key factors in establishing confidence in teachers themselves and of others in teachers' judgements.
- There needs to be clarity about the operational details of systems based on teachers' judgements, followed by at least two years of trial and evaluation for their value to be properly judged.
- Interactions among curriculum pedagogy and assessment have to be taken into account so that changes in any one are reflected in the others; for instance, the widely advocated broadening of the curriculum is reflected in a greater range of approaches to assessment.

## **Appendix 2 Key findings from systematic reviews of research related to teachers' summative assessment**

Two linked systematic reviews (EPPI reviews) reported evidence from research relating to

1. The reliability and validity of assessment by teachers for summative purposes and the conditions that affect its reliability and validity.
2. The impact on students, teachers and the curriculum of the process of using assessment by teachers for summative purposes and the conditions and contexts that affect the nature and extent of any impact.

*Main findings from the first review relevant to the subject of this paper were:*

1. The accuracy, or reliability, of teachers' assessment can be increased by increasing the degree of specification of the task and/or of the criteria used in judging the process or outcome.
2. Greater specification of tasks tends to narrow the range of evidence used and so lowers its validity. The alternative of finer specification of criteria, particularly describing progressive levels of competence, has been shown to be capable of supporting reliable TA whilst allowing evidence to be used from the full range of classroom work.
3. The clearer teachers are about the goals of students' work, the more consistently they apply assessment criteria.
4. There is evidence from comparing TA with test scores suggesting some bias in teachers' judgements due to teachers taking account of characteristics of students not related to the assessment criteria, such as behaviour (in the case of young children), gender, special educational needs, verbal ability and overall academic achievement.
5. Training can address bias and improve the reliability of TA particularly when teachers participate in developing criteria, have some ownership of them and understand the language used.
6. Moderation through professional collaboration is of benefit to teaching and learning as well as to assessment. There needs to be protected time for teachers to meet as a regular part of their professional commitments.

*Main findings from the second of these reviews were:*

7. Teachers' summative assessment practices are likely to have a positive impact on teaching if teachers find them of value in helping them to learn more about their students and to develop their understanding of curriculum goals. Time to experience and develop some ownership of new practices enhances their positive impact.
8. When high stakes judgements are associated with teachers' assessment, one effect is for teachers to reduce assessment tasks to routine events and restrict students' opportunities for learning from them. High stakes encourages some teachers to give high grades where there is doubt, which may not be in the students' interests.
9. Teachers using shared criteria relating to specific aspects of achievement has a positive impact on students and on teaching; in the absence of such guidance there is little positive impact on teaching and a potential negative impact on students.
10. The process that teachers use in setting assessment tasks and in grading, impacts on students' motivation for learning, particularly their goal orientation, when grades are used as rewards or punishments. The negative impact can be alleviated by ensuring that students have a firm understanding of assessment processes and criteria
11. Summative assessment by teachers has a more positive impact on teachers and teaching when integrated into practice than when concentrated at a certain occasions.
12. Opportunities for teachers to share and develop their understanding of assessment procedures enable them to review their teaching practice, their view of students' learning and their understanding of subject goals. Such opportunities require to be sustained over time and preferably should include provision for teachers to work collaboratively across as well as within schools.

---