

- Students should participate in setting targets, and have a sense of ownership of the targets. (Tanner and Jones 2003c, p. 69).

A formal process of target-setting that fulfilled these criteria would be very time consuming, and while it would be possible to prioritize the process at the expense of other activity, the ambiguous nature of the research evidence makes us doubt that this is the best use of resources.

We do not doubt, however, the value of asking children to reflect on and evaluate their work, identifying their own learning needs. Nor do we doubt the value of teachers providing regular feedback that offers children advice on how to improve their learning. But such activities should be informal in character and part of the normal day-to-day feedback and marking strategies that are planned to be a part of a systematic process of formative assessment.

Planning for assessment in the short, medium and long term

As we indicated in the last chapter, children should be asked to pause to reflect on and evaluate their learning on a regular basis. Reflection is supposed to underpin the target-setting process and, in the best practice, target-setting meetings are reflective. However, reflection alone is not enough – action-planning should result.

When target-setting is based on long-term aspirations and performance rather than learning, the danger is that targets will be more like New Year resolutions rather than specific action plans. If target-setting is to stimulate intellectual development, then learning tasks, assessment techniques and feedback must all be planned to challenge misconceptions, support reflection and help students to identify clear steps towards the re-structuring of ideas.

You should realize by now that assessment is not something that can just happen in a haphazard way, but should be planned for. Indeed it is as important to plan assessment as it is to plan teaching. They are opposite sides of the same coin and should be planned together.

In order to be able to teach and assess effectively, you need to be clear about your aims. By this we mean more than giving clear learning outcomes for teaching episodes. It is easy to fall into the trap of reducing your subject to a list of detailed learning objectives that can be assessed easily, but to lose sight of your long-term aims for teaching your subject. The assessment regime in

your classroom ought to reflect your beliefs about learning and teaching and the nature of your subject.

The prevailing assessment regime has a major influence on the development of the attitudes and norms of behaviour that form the basis of the learning culture in your classroom. Many powerful cognitive and emotional messages are conveyed through the ways that you assess your students and you should think very carefully about the messages that you are sending. Assessment provides information, but for whom? Is it aimed at providing feedback information to students so that they can modify their learning, or is it aimed at collecting data for records and reporting to others? Is assessment done *for* students or *to* students?

It is not possible to formally assess everything that a student does. You have to select key skills, knowledge and concepts as indicators of learning. At a very practical level, you have to decide *what* you assess, *when* you assess it and *how* you assess it. If the learning outcomes that you have planned for your lessons are sufficiently rich, they should lead you naturally to what you wish to assess. However, WYTIWYG – What You Test Is What You Get – is a helpful acronym to bear in mind when determining what to assess. What you choose to assess sends a powerful message to your class about what you value in your subject. Although you may claim to value creativity and imagination, if your assessments reward only memorization, your class will draw the obvious conclusions. We must ensure that we assess what is important rather than that which is convenient.

There is a tension here, as teachers often try to write assessment items in the style of external examination questions. Because of the high stakes nature of external examinations, these aim at high reliability rather than validity and are sometimes quite limited in scope. If your assessment is intended to serve formative aims rather than predict examination performance, you may choose to emphasize validity over reliability and provide feedback on a wider range of knowledge and skills. The dead

hand of external examination should not be allowed to permeate the whole of secondary education.

When and how you assess will also depend on the precise purpose of the assessment, and whether you wish to record, predict or modify performance. If your purpose is to record performance, you should be clear about the purpose (if any) of the record. Unless reinforced by review or revision, knowledge appears to decay over time. Recorded data from previous assessments may offer only limited information about current understanding or future performance. In spite of the effort expended on them, many of the records maintained in school seem unfit for the purposes they are intended to serve (Black and Wiliam 1998a, b).

The timing, context and form of an assessment make a significant difference to its usefulness. When planning for assessment, we find it helpful to consider these issues in the short, medium and longer term.

Short-term formative assessment is dynamic and flexible in character, designed to initiate immediate, real-time changes to learning and teaching. The results are short-lived and there is little point in recording them – they should be out of date before the end of the lesson. These dynamic short-term assessments fall into three broad categories: baseline, concurrent and plenary.

Informal baseline assessments usually occur at the start of the lesson as you set the scene, confirming prior knowledge and identifying misconceptions. As the lesson proceeds, concurrent formative assessment is used to fine-tune interactions, helping you to judge the minimum amount of support needed by individuals. Skilled teachers also use concurrent assessment to identify specific misconceptions that they then target with supplementary questions.

This process is flexible and dynamic, and sometimes appears to be unplanned, because there is no pre-prepared script and the exact route that the lesson will follow is not fixed in advance. Certainly there is an element of 'flying by the seat of your pants'

inherent in the process when experienced teachers make themselves responsive to children's thinking. Trainee teachers sometimes mistakenly read this as suggesting that the most effective and experienced teachers do not plan their lessons. This is not the case. We refer to the form of planning involved as 'fuzzy planning'. Learning objectives and potential misconceptions are identified clearly in advance. Key 'rich questions' are selected to challenge misconceptions. Resources are prepared and pivotal points in the lesson are identified. The general trajectory of the lesson is known even if the precise path is not.

Further formative assessment should occur in the plenary. The plenary provides an opportunity for students to pause and reflect on their learning, consolidating and formalizing the knowledge they have gained and evaluating the extent to which they have achieved the learning objectives. This should also be a point where areas of unease or incomplete understanding are identified and personal learning targets are formulated. There are several plenary activities that can support such reflective formative assessment, and we will discuss these further in Chapter 4: however, most are dependent on effective feedback and support by the teacher in the form of rich questioning to stimulate discussion and the co-construction of knowledge. An effective plenary session also provides feedback to the teacher, allowing you to evaluate the success of your teaching in relation to the learning of individual students, and to plan for the next lesson.

Information for medium-term assessments may potentially come from a number of different sources. These might include: tasks completed in class or at home, marking exercise books, informal short-class tests, or more formal end-of-unit/half term/end-of-term tests. Your school or department probably has an assessment policy that details which of these sources are to be used, and the minimum frequency of assessment.

To use end-of-unit assessments formatively demands careful timing. Although, at first sight, it seems logical to assess at the

end of the module, this is not helpful for formative purposes. By the time the feedback is received by the student that unit is finished and new work is being studied. There is little for the student to do but note the grade and resolve to try harder next time. To serve formative ends, assessment and feedback should occur before the end of the module, allowing students to remedy weaknesses and improve their performance while there is still time. As a student, the focus then moves away from assessment being done to you to provide grades and towards providing feedback to you to help you learn.

If the form of the assessment is more imaginative than a timed written test, perhaps involving making a presentation to the class, or writing a report on a short investigation, then formative feedback may be given via discussion of a first draft prior to final submission at the end of the module.

Medium-term assessments are commonly used to generate grades to monitor or track students' progress. The grades are often used when reporting to parents, providing evidence, for example, that a child is good at algebra but weak at geometry. The extent that such information is actually useful is at best debatable. It is not clear how the child or the parents are supposed to use such information to effect change. More finely grained data, based on detailed analysis of strategies and errors, would be required to be of use in setting effective learning targets. Unfortunately, the analysis of students' work to determine learning needs is given a low priority in most schools (Black and Wiliam 1998b, p. 6).

Some schools store such data centrally and use it to monitor or track students so that they can identify trends in performance, perhaps identifying that a child has begun to underperform across a number of subjects. Systems like this are intended to prevent students from 'falling through the net' by identifying potential problems in good time. However, it is not the records themselves that are significant, but the use that is made of them within the pastoral system.

The value of end-of-unit assessment records for predicting performance in external examination is often dubious. An assessment made immediately after teaching is likely to overestimate performance at a later date, particularly if the topic has not been revisited since. Information that has been learned in a superficial manner may be retained in the short term, but decay rapidly in the weeks that follow. Records that are to be used for prediction require an assessment of deeper, more meaningful learning. Furthermore, examining one unit at a time fails to assess whether students have made links between different topics and are able to apply knowledge in different contexts. End-of-unit tests appear to be the wrong mechanism for the purpose of prediction. Although medium-term assessments based on topics or units of work have a minor summative function, the emphasis should be placed on their formative functions and they should be timed accordingly.

Summative and predictive aims would be better served by intermittent assessments on a longer timescale. Medium-term summative assessments should include at least two or three new topics, and also periodically reassess significant key objectives from older work. Such a system could be used to encourage students to review older work periodically. Much depends on the preparation that precedes the assessment, the culture of the classroom and the extent to which revision skills are taught directly.

Currently, the primary aims of long-term assessments are often managerial in character. Students' progress may be measured against national standards or the key objectives for a key stage, allowing managers of the system to hold teachers, departments or schools accountable, or to judge the impact of an initiative. Value-added data are often used, allowing managers to monitor the performance of classes and teachers towards predicted targets.

Secondary schools usually undertake a summative assessment of their students at least once a year. These assessments are often based on the results of school tests or examinations, which summarize the work of the year or the term. These examinations are

often very formal, time consuming, disrupt the timetable, and represent a significant event in students' school lives.

It is often claimed that end-of-year summative assessments provide useful information about standards for the next teachers of the students. Sadly, the information provided by end-of-year examinations is rarely in a form that would aid the planning of a new teacher taking over the class.

The precise purpose of end-of-year examinations is not always clear, beyond providing an apparently objective item to include on a school report! They may also be used to justify allocating a student to a particular ability set or course. However, you might question whether the time and emotion invested in such examinations is worth the effort for such limited purposes. Surely good teachers should know their students well enough to discuss their progress with parents or to judge if they are misplaced in a class?

When school examinations have high status, they can motivate some children to revise work, although few know how to do this effectively without significant intervention from teachers. There is also evidence that when the focus is on competition and comparisons between students, there is a negative impact on the development of some groups of students. While competition can be a powerful motivator, the losers in the competition often begin to believe that they lack natural ability, leading them to 'retire hurt', losing the motivation to learn (Black 1998b, p. 43). Much depends on whether the assessment culture in the school is focused on performance or learning.

We are not arguing that long-term summative assessment is unimportant – far from it! Assessment by single topic is insufficiently challenging and fails to encourage students to consolidate learning from different topics over time. Furthermore, in order to maintain a sense of direction and purpose, learners need to stand back and take stock at regular intervals. However, many secondary school examination and recording systems still suffer from an absence of educational purpose. The collection of marks to fill

up records is often given a higher priority than the analysis of students' work to discern learning needs (Black and Wiliam 1998b, p. 6). Ofsted (1998, p. 93) reported that although the recording of student progress was improving, the results were generally underused, which suggests that sadly in many schools, summative assessment and recording remains an end in itself rather than an aid to learning.

Although the primary purpose of longer term assessment is often seen as summative, it is possible for longer term assessments and examinations to serve formative purposes. Much depends on the revision and note-taking skills that children are taught to use in preparation beforehand, and the analysis of results and feedback that occurs after the examination. We will discuss the teaching of revision strategies and the use of examinations to support learning in Chapter 5.

Longer term assessments such as yearly examinations should represent the culmination of an assessment and learning strategy based on short- and medium-term formative assessments. The form and timing of such assessments and the feedback provided contribute significantly to the development of the learning culture within your school, department or classroom. If assessment is to support learning and produce the gains described in the research literature, then the culture should emphasize learning rather than performance; personal improvement rather than competition; and developmental rather than innate ability (Black and Wiliam 1998a, b). Much of this depends on the form of feedback offered to students when assessing their written work.

What kind of feedback do students need?

Although significant feedback occurs in real time in the two-way process of interactive teaching, inevitably much individual feedback is based on the assessment and monitoring of students'

written tasks, exercises, tests and examinations. Although some feedback may be given orally during lessons while the work is progressing, inevitably much of the assessment of children's work occurs outside the classroom, and the feedback is often provided in written form, through marking.

If marking and its associated feedback is not planned effectively, it can easily lose its formative qualities and degenerate into checking whether work has been completed and presented well, or the checking of low-level facts. In fact, analysis of 131 research studies on feedback by Kluger and DeNisi (1996) showed that although on average there was a strong positive impact, in roughly 40 per cent of studies, the effect was negative. The form of feedback, the use made of it and the culture of the classroom makes a significant difference.

Black and Wiliam (1998a) claim that if assessment is to serve formative purposes, then certain key components must be present in the feedback:

- a clear indication of the nature of a perfect answer;
- information about the standard actually achieved;
- advice about how the gap might be closed;
- monitoring of the student's response to the advice in future work. (Adapted from Sadler 1989, p. 121; Black and Wiliam 1998a, p. 48).

Your interpretation of their first criterion is likely to depend on the nature of your subject and the task that you are imagining. In many closed tasks, the nature of a perfect answer may be relatively unproblematic for a student to interpret. However, in more open-ended tasks, such as coursework investigations, determining the nature of a good solution may be a significant issue. The second and third components are expressed in terms that suggest externally imposed benchmarks or examinations rather than learning