

## Chapter 6

### The role of formative assessment in effective learning environments

**Dylan Wiliam**

Institute of Education, University of London

*Dylan Wiliam describes assessment as the bridge between teaching and learning. The concept of “formative assessment” emerged with recognition of the importance of feedback and application of navigational metaphors about staying on course through corrective steering. There is substantial evidence, reviewed here, on how feedback improves learning but most studies suffer from weak conceptualisation and neglect of longer-term impacts. The definition here emphasises the role of assessment in improving the quality of instructional decisions. It can be seen as entailing five “key strategies”:*

- 1. Clarifying, sharing and understanding learning intentions and criteria for success.*
- 2. Engineering classroom activities that elicit evidence of learning.*
- 3. Providing feedback that moves learners forward.*
- 4. Activating students as instructional resources for one another.*
- 5. Activating students as owners of their own learning.*

*Formative assessment is proposed as a process of capitalising on, “moments of contingency” for the purpose of regulating learning processes.*

## Introduction

Assessment plays a number of roles in modern societies, including the certification of student achievement and holding educational institutions to account. Over the past approximately 40 years, however, there has also been increasing interest in the role it can play in supporting learning, often called “formative assessment” or “assessment for learning”. This chapter presents a brief overview of how the concept of formative assessment has developed in recent years; in particular, how the central idea has expanded from an original focus on feedback to a wider perspective on classroom practice. It presents evidence on the impact of formative assessment on learning and discusses definitional issues. It concludes with discussion on how formative assessment relates to instructional design through the “regulation” of learning processes.

## Why assessment is central to learning

If what students will learn as the result of a particular sequence of activities were predictable, designing learning would be simple. Provided that we ascertain that students possess the correct prerequisites for a particular learning sequence, we could be sure that they all would have learned what was intended after engaging in the specified activities. However, as Denvir and Brown found (1986a; 1986b), even when teachers design high quality learning activities aimed at particular skills, and even when they take into account the student’s prior knowledge, what is learned can often be quite different from the intended goal.

Yet, in most classrooms across the world, evidence about the success of learning activities is typically collected only at the end of the learning sequence. It is as if the crew of an aircraft on a long journey concentrated only on following the optimal course from their starting point to their destination, and paid no attention to whether they were, in fact, on course. As all pilots know, this is an unreliable strategy. This is why, in addition to plotting a careful course, aircrew also take readings of their position as they are heading towards their destination and make adjustments as conditions dictate.

In a similar vein over 40 years ago, Benjamin Bloom suggested that in addition to assessment used at the end of a learning process to establish what had been learned, assessment could also be used “to provide feedback and correctives at each stage in the teaching-learning process” (Bloom, 1969 p. 48). He also noted that, while such assessments “may be graded and used as part of the judging and classificatory function”, it is much more effective “if it is separated from the grading process and used primarily as an aid to teaching” (p. 48).

David Ausubel stated many years ago: “If I could reduce all of educational psychology to one principle, I would say this: the most important single factor influencing learning is what the learner already knows. Ascertain this and

teach him accordingly” (Ausubel, 1968 p. iv). Assessment is central to effective learning, therefore, because even if learners start in roughly the same place with respect to a particular piece of learning, they will very quickly be at different places due to the differences in what they have learned.

This is the fundamental idea explored in this chapter: the design of learning environments needs to take account of the fact that learning is unpredictable so that assessment has a key role to play by relating the instructional activities that teachers plan to the consequent increase in learner capabilities. In other words, assessment functions as the bridge between teaching and learning. The aim of this chapter is to provide a clear theoretical basis for the ways in which assessment can support learning, to show how the different formulations of the notion of formative assessment proposed over the last 40 years can be encompassed within a broader over-arching framework, and to use that framework to understand research in related areas.

### **Formative assessment as feedback**

Course correction in navigation as discussed above is an example of a “feedback” system, developed originally in the field of systems engineering (see Wiener, 1948). Wiener noted that sometimes the effect of the “feedback loop” is to drive the system further in the direction it is already going, such as population growth with plentiful food and no predators or inflationary price/wage spirals in economics. Such feedback is called “positive feedback” because the effect of the feedback and the tendency of the system operate in the same direction. In other situations, the effect of the feedback is to oppose the tendency, restoring stability by returning the system to a steady state, as with population growth when food supply is limited or the familiar room thermostat. This is called “negative feedback” by engineers since its effect is in the opposite direction to the tendency of the system. In engineering, positive feedback is unhelpful because it means instability leading either to explosive growth or collapse. In contrast, negative feedback helps to restore the system to a stable state.

The metaphor of “feedback” is widespread in education but it is important to note that there are significant differences between the usage of the term in engineering and in education. First, to qualify as feedback for an engineer, the system must be able to use the information to affect its performance: “Feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way.” (Ramaprasad, 1983, p. 4) In contrast, in education the term “feedback” is often used to describe any information given back to a learner about their performance, irrespective of whether that information has the capacity to alter the gap (Sadler, 1989). In other words, if we use the term as an engineer would, feedback is not just information given to students about their performance. It must direct their future actions in productive ways.

Second, not only the term “feedback” but the qualifiers “positive” and “negative” are also applied in somewhat different ways. In engineering, they refer to the effect of the feedback in relation to the tendency of the system. In education, the terms tend to be used instead as value judgments on the effects of the feedback. Feedback that suggests that the learner is on the right track, so reinforcing the learning, would be described as “positive” both by educators and engineers. However, consider the situation in which a student received critical evaluations, made less effort, got even worse evaluations and made even less effort, ultimately disengaging from learning altogether. To an educator, this is an example of negative feedback but to an engineer this is positive feedback, since it drives the system (student) in the direction it is already heading.

Third, and perhaps most importantly, we want in education to encourage the development of autonomy in learning – for students to be able to develop their own skills of self-regulation of learning so that their need for feedback diminishes. In contrast, no-one would criticise a room thermostat because the furnace had not yet learned when to decide for itself when to turn itself on and off.

While these may appear to be semantic distinctions, in fact they go to the heart of the problems encountered in the design of effective feedback systems in education. Crooks (1988) reviewed over two hundred studies of the impact of classroom evaluation practices on students and concluded that the power of assessments to guide learning was not being realised because the summative function of assessment – providing grades and other measures of how much had been learned – is dominant.

### ***Evidence on the impact of feedback***

Studies have found that feedback can substantially improve educational outcomes but we should be aware of certain caveats by way of introduction. The results of many studies are given in terms of a “standardised effect size” [“effect size” for short: this following Cohen (1988) is the difference in performance between two groups (*e.g.* those given and those not given feedback) divided by a measure of the spread of scores in the population (the standard deviation)]. While the standardised effect size has undoubted advantages over reporting the level of statistical significance attained in experimental comparisons (Harlow, Mulaik and Steiger, 1997), it nevertheless suffers from limitations as a metric with which to compare findings from different experimental studies. In particular, where the range of outcomes is restricted (*e.g.* studies on specific sub-populations such as students with special educational needs), the effect size is inflated because the divisor in the calculation is smaller (Black and Wiliam, 1998a). Second, measures of educational outcomes differ greatly in their sensitivity to the effects of education and whether the measure relates directly to what students have been learning or is more remote, as with many national tests and examinations (Wiliam, 2008). This means that

it is difficult to give hard-and-fast rules about how to interpret effect sizes. Nevertheless, as a general guide, at least on standardised measures of educational achievement, effect sizes of around 0.4, which are typical in studies of feedback, indicate an increase of at least 50% in the rate of learning. In other words, students were learning in 8 months what other students were taking a year to learn. These are therefore rather substantial increases in educational productivity, especially if they can be scaled across an entire national system.

A more general caveat is that evaluations are used in schools for a multiplicity of purposes and comparisons are misleading when evaluations are compared in terms of functions for which they were not designed (e.g. Natriello, 1987). For example, finding that differentiated feedback has more impact on directing future student learning than on grades may show nothing more than that systems generally do more effectively those things they are designed to do than those things they are not designed to do.

Such limitations notwithstanding, the first substantial finding is that just being assessed regularly can have a significant impact on learning. For instance, students who took at least one test over a 15-week period scored 0.5 standard deviations higher than those who did not, and more frequent testing was associated with higher levels of achievement, although testing more frequently than once every two weeks conferred no additional benefit (Bangert Drowns, Kulik, Kulik and Morgan, 1991). The quality of feedback and how it is used, however, are much more important than its frequency. A review of 40 research reports on the effects of feedback in “test-like” events (such as questions embedded in programmed learning materials or review tests at the end of a block of teaching) found that the way feedback was provided and the kind of feedback given were both critical (Bangert-Drowns *et al.*, 1991). Where students could look ahead and “peek” at the answers before they had attempted the questions, they learned less than when studies controlled for this “pre-search availability” (effect size: 0.26). More importantly, when feedback is given through the details of the correct answer, students learn more than when they are just told whether their answer is correct or not (effect size: 0.58).

Feedback can also be useful to teachers. Fuchs and Fuchs (1986) conducted a meta-analysis of 21 different reports on the use of the feedback to and by teachers, with frequencies of between 2 and 5 times per week. The mean effect size on achievement between experimental and control groups was 0.70 standard deviations. In about half the studies reviewed, teachers set rules about reviews of the data and actions to follow and in these cases the mean effect size was significantly higher at 0.92; when actions were left to teachers’ judgments the effect size was only 0.42. In those studies in which teachers produced graphs of the progress of individual children as a guide and stimulus to action, the effect was larger (mean effect size: 0.70) than in those where this was not done (mean effect size: 0.26).

These findings appear to be affected by the kind of learning being considered. Dempster (1991) found that many of the available research studies measured achievement in terms of content knowledge and low-level skills so that it is not clear that such findings would necessarily generalize to higher-order thinking. In a subsequent paper, Dempster (1992) argued that while the benefits of integrating assessment with instruction are clear, and there is an emerging consensus in the research for the conditions for effective assessment – frequent testing soon after instruction, cumulating demand, with feedback soon after testing – assessment is neglected in teacher education and current practices in schools are far from these ideals.

A review by Elshout-Mohr (1994), published originally in Dutch and reviewing many studies not available in English, suggested that for more complex tasks, knowledge of correct answers is less useful than it is for simple tasks. Learning is not just a matter of correcting what is wrong but of developing new capabilities and this requires feedback more as dialogue rather than simply giving correct answers. This requires the learner to become active in managing the process.

Much of this work had focused on the effects of feedback in schools. In 1996, Kluger and DeNisi published a review of the effects of feedback in schools, colleges and workplaces.<sup>1</sup> Across all the studies, the average effect size for the feedback is 0.41 standard deviations, but the effects vary considerably across the different studies. Most notably in 50 out of the 131 studies (38%) feedback actually **lowered** average performance.

As part of a broader research programme on the development of intelligent tutoring environments, Shute (2008) examined research on feedback to students.<sup>2</sup> This review identified major gaps in the literature and, as might be expected, concluded that there was no simple answer to the question, “What feedback works?”. But, it also endorsed the findings of earlier reviews on the size of the effects that could be expected from feedback (standardized effect sizes in the range 0.4 to 0.8 standard deviations).

### *Some pointers regarding effective feedback*

In seeking to understand why feedback may sometimes lower performance, Kluger and DeNisi (1996) looked for “moderators” of feedback effects. They found that feedback was least effective when it focused attention on the self, more effective when it focused on the task in hand, and most effective when it focused on the details of the task and involved goal-setting.

However, even the limited benefits of feedback identified by Kluger and DeNisi might sometimes be counter-productive. They pointed out that feedback might make the learner work harder, which is presumably beneficial, but it might also lead the learner to channel her or his efforts in a particular direction,

to modify or reject the goal, or to ignore the feedback entirely. Even when feedback produced a positive impact on learning, this might be by emphasising instrumental goals and inhibiting deep learning. In their conclusion, they suggested that it is more important to examine the processes induced by the feedback rather than whether feedback in general improves performance.

Shute (2008) offers a number of “preliminary guidelines” for the design of effective feedback, both in relation to enhancing learning and in terms of timing.

Feedback should focus on the specific features of the task, and provide suggestions on how to improve, rather than focus on the learner; it should focus on the “what, how and why” of a problem rather than simply indicating to students whether they were correct or not; elaborated feedback should be presented in manageable units and, echoing Einstein’s famous dictum, should be “as simple as possible but no simpler.” However, feedback should not be so detailed and specific that it “scaffolds” the learning to such an extent that the students do not need to think for themselves. Feedback is also more effective when from a trusted source (whether human or computer).

The optimum timing of feedback appears to depend strongly on the kind of learning being undertaken: immediate feedback appears to be most helpful for procedural learning or when the task is well beyond the learner’s capability at the beginning of the learning, while delayed feedback appears to be more appropriate for tasks well within the learner’s capability or when transfer to other contexts is sought.

The recent review by Hattie and Timperley (2007) defines the purpose of feedback as reducing discrepancies between current understandings or performance and a desired goal (as proposed by Ramaprasad, 1983). Building on the work of Deci and Ryan (1994) and Kluger and DeNisi (1996), their model posits that students can reduce the discrepancy either by employing more effective strategies or increasing effort, on the one hand, or by abandoning, blurring or lowering the goals they have set for themselves, on the other. Teachers can reduce the discrepancy by changing the difficulty or specificity of the goals or by providing more support to the students. Their model specifies three kinds of questions that feedback is designed to answer (Where am I going? How am I going? Where next?), and each feedback question operates at four levels: feedback about the task (FT), feedback about the processing of the task (FP), feedback about self-regulation (FR), and feedback about the self as a person (FS). They demonstrate that FS is the least effective form of feedback; FR and FP “are powerful in terms of deep processing and mastery of tasks”; FT is powerful when the feedback is used either to improve strategy processing, or for enhancing self-regulation (although these conditions are rarely met in practice).

## Formative assessment as part of teaching

The studies summarised above show that some form of feedback to learners in the course of their learning has positive effects on learning, but that such effects cannot be taken for granted. The effects depend not just on the quality of the feedback but on the learning milieu in which it is provided, the orientations and motivations of the learner, and a range of other contextual factors (Boekaerts, this volume). For this reason, when Paul Black and I sought to update the reviews of Natriello and Crooks, we deliberately took a broad view of the field. (We noted that the reviews by Natriello and Crooks had cited 91 and 241 references respectively, and yet only 9 references were common to both papers, and neither cited the review by Fuchs and Fuchs.) Rather than relying on electronic search methods, we consulted each issue of 76 of the journals considered most likely to contain relevant research between 1987 and 1997. Our review (Black and Wiliam, 1998a), based on 250 studies, found that effective use of classroom assessment yielded improvements in student achievement between 0.4 and 0.7 standard deviations, albeit noting the already-mentioned problems with the interpretation of effect sizes.

Black and Wiliam presented a number of “examples in evidence” – the meta-analysis by Fuchs and Fuchs and seven classroom-based studies – that illustrate features of effective formative assessment. Perhaps the most important one is that, to be effective, formative assessment has to be integrated into classroom practice, requiring a fundamental re-organisation of classroom operations:

It is hard to see how any innovation in formative assessment can be treated as a marginal change in classroom work. All such work involves some degree of feedback between those taught and the teacher, and this is entailed in the quality of their interactions which is at the heart of pedagogy. (Black and Wiliam, 1998a, p. 16)

We also noted that for assessment to function formatively, the feedback information has to be used, and thus the differential treatments that are incorporated in response to the feedback are at the heart of effective learning. Moreover, for these differentiated treatments to be selected appropriately, teachers need adequate models of how students might react to, and make use of, the feedback. As Perrenoud (1998) observes in his commentary on the Black and Wiliam paper, “...the feedback given to pupils in class is like so many bottles thrown into the sea. No one can be sure that the message they contain will one day find a receiver.”

In order to address this, we examined the **student** perspective, the role of **teachers**, and some of the **systems** for the organisation of teaching in which formative assessment is a major component. In drawing out implications for the policy and practice of formative assessment, we concluded:



There does not emerge, from this present review, any one optimum model on which ... policy might be based. What does emerge is a set of guiding principles, with the general caveat that the changes in classroom practice that are needed are central rather than marginal, and have to be incorporated by each teacher into his or her practice in his or her own way .... That is to say, reform in this dimension will inevitably take a long time and need continuing support from both practitioners and researchers. (p. 62)

Most of the work reviewed by Natriello, Crooks, Kulik and his colleagues, and Black and Wiliam focused on school-age students (*i.e.* up to the age of 18). Nyquist (2003) examined studies of feedback with college-age learners. He reviewed approximately 3000 studies of the effects of feedback, of which 86 met the criteria that they:

- Involved experimental manipulation of a characteristic relevant to feedback.
- Used a sample of college-age learners.
- Measured academic performance.
- Provided sufficient quantitative information for an effect size to be calculated.

From the 86 studies it was possible to derive 185 effect sizes. After a number of technical adjustments (limiting extreme values to 2 standard deviations from the mean effect, and correcting for small sample bias across the studies), the analysis yielded a mean effect size of 0.40 standard deviations – almost identical to that found by Kluger and DeNisi. This mean effect reduced slightly to 0.35 (SE = 0.17) once adjustments were made (weighting the effects so that the contribution to the mean effect was proportional to their reliability), although the effects themselves were highly variable (ranging from -0.6 to 1.6 SDs).

To investigate “moderators” of effect, Nyquist developed the following typology of different kinds of formative assessment:

- **Weaker feedback only:** students are given only the knowledge of their own score or grade; often described as “knowledge of results”.
- **Feedback only:** students are given their own score or grade, together with either clear goals to work towards or feedback on the correct answers to the questions they attempted; often described as “knowledge of correct results”.
- **Weak formative assessment:** students are given information about the correct results, together with some explanation.

- **Moderate formative assessment:** students are given information about the correct results, some explanation, and some specific suggestions for improvement.
- **Strong formative assessment:** students are given information about the correct results, some explanation, and specific activities to undertake in order to improve.

The average standardized effect size for each type of intervention is given in Table 6.1.

Table 6.1. **Effect sizes for different kinds of feedback intervention**

	N	Effect
Weaker feedback only	31	0.14
Feedback only	48	0.36
Weaker formative assessment	49	0.26
Moderate formative assessment	41	0.39
Strong formative assessment	16	0.56
Total	185	

*Source:* Nyquist, 2003. The figures are corrected values provided in a personal communication and not the same as given in the original thesis.

Nyquist's results echo the findings of Bangert-Drowns *et al.* discussed above. Just giving students feedback about current achievement produces relatively little benefit, but where feedback engages students in mindful activity, the effects on learning can be profound.

The research reviews conducted by Natriello (1987), Crooks (1988), Bangert-Drowns *et al.* (1991), and Black and Wiliam (1998a) underline that not all kinds of feedback to students about their work are equally effective. As a further example, Meisels, Atkins-Burnett, Xue, Bickel and Son (2003) explored the impact of the Work Sample System (WSS) – a system of curriculum-embedded performance assessments – and the achievement of WSS students was significantly and substantially higher in reading, but in mathematics there was no significant difference. The details of the system in use, how it is implemented, and the nature of the feedback provided to students appear to be crucial variables, with small changes often producing large impacts on effectiveness.

Though many of the studies included in the reviews focus on older students, attitudes to learning are shaped by the feedback they receive from a very early age. In a year-long study of eight kindergarten and first grade

classrooms in six schools in England, Tunstall and Gipps (1996a; 1996b) identified a range of roles played by feedback. Like Torrance and Pryor (1998), they found that much of the feedback given by teachers to students focused on socialisation: “I’m only helping people who are sitting down with their hands up” (p. 395). Beyond this socialisation role, they identified four types of feedback on academic work.

The first two types are essentially **evaluative** in form. The first covers feedback that rewards or punishes the students for their work (e.g. students being allowed to leave for lunch early when they had done good work, or threatened with not being allowed to leave for lunch if they hadn’t completed assigned tasks). The second type of feedback is also evaluative, but indicates the teacher’s level of approval (e.g. “I’m very pleased with you” vs. “I’m very disappointed in you today”). The two other types of feedback identified by Tunstall and Gipps are termed “descriptive”. The third focuses on the adequacy of the work in terms of the teacher’s criteria for success, ranging from the extent to which the work already satisfies the criteria at one end (e.g. “This is extremely well explained”) to the steps the student needs to take to improve (e.g. “I want you to go over all of them and write your equals sign in each one”). The fourth kind of feedback emphasises process, with the teacher playing the role of facilitator rather than evaluator. As Tunstall and Gipps (1996a) explain, teachers engaging in this kind of feedback “conveyed a sense of work in progress, heightening awareness of what was being undertaken and reflecting on it” (p. 399).

Most of the research reviewed above was published in English. In order to provide a more comprehensive overview of research in this area, the OECD study of formative assessment (Looney, 2005) commissioned reviews of relevant research published in French (Allal and Lopez, 2005) and German (Köller, 2005).

Allal and Lopez report that research in France and French-speaking parts of Belgium, Canada and Switzerland, has focused much more on theoretical than empirical work, with very few controlled empirical studies. They suggest that the most important finding of their review of over 100 studies of the previous thirty years is that the studies of assessment practices in French-speaking classrooms have utilized an “enlarged conception of formative assessment” along the lines adopted by Black and Wiliam. Allal and Lopez argue that central to feedback within the Anglophone tradition (as exemplified by Bloom), is “remediation,” which they summarise as “feedback + correction”. In contrast, within much of the research undertaken in francophone countries, the central concept is “regulation”, summarised as “feedback + adaptation” (p. 245).<sup>3</sup>

Allal and Lopez identify four major developments in this French-language research literature. In the first, which they term “focus on instrumentation”,

the emphasis was on the development of assessment tools such as banks of diagnostic items and adaptive testing systems. In the second (“search for theoretical frameworks”), the emphasis shifted to a “search for theories that can offer conceptual orientation for conducting assessment”. The third development – “studies of existing assessment practices in their contexts” – provides a grounding for the search for theoretical frameworks by articulating it with the study of how formative assessment is practised in real classrooms. The fourth, and most recent, development has been “development of active student involvement in assessment” which has examined student self-assessment, peer-assessment, and the joint construction of assessment by students and teachers together.

The notion of formative assessment as being central to the regulation of learning processes has been adopted by some Anglophone researchers (see, for example, Wiliam, 2007), and the broadening of the understanding of formative assessment was noted by Brookhart (2007). Her review of the literature on “formative classroom assessment” charted the development of the concept of formative assessment as a series of nested formulations (p. 44):

- Information about the learning process.
- Information about the learning process that teachers can use for instructional decisions.
- Information about the learning process that teachers can use for instructional decisions and students can use in improving their performance.
- Information about the learning process that teachers can use for instructional decisions and students can use in improving their performance in ways that motivate them.

In general, however, there appear to be few links between the strong theoretical work in the francophone tradition and the empirical work undertaken particularly in the United States. Allal and Lopez conclude that the French-language work on formative assessment is in need of considerably more empirical grounding. (p.256)

The review of German-language literature by Köller (2005) began with an approach similar to that adopted by Black and Wiliam, with searches of on-line databases supplemented by scrutiny of all issues of the six most relevant German-language journals from 1980 to 2003. Köller noted that while there were many developments related to formative assessment reported in academic journals, there was little evaluation of the outcomes of formative assessment practices for students, although there were confirmations of some findings in the Anglophone literature. He reports the work of Meyer who, like Kluger and DeNisi, found that praise can sometimes have a negative impact on learning, while criticism, even blame, can sometimes be helpful. Another

important strand of work mentioned by Köller concerns differences between teachers' uses of "reference norms." A number of studies, notably those by Rheinberg, have shown that students learn more when taught by teachers who judge a student's performance against his or her previous performance (individual reference norm) rather than teachers who compare students with others in the class (social reference norm).

### Theoretical syntheses: formative assessment and assessment for learning

Over the last dozen or so years, a number of definitions of the term "formative assessment" have been proposed. Black and Wiliam (1998a) defined formative assessment "as encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" (p. 7). Cowie and Bell (1999) adopted a slightly more restrictive definition by limiting the term to assessment conducted **and acted upon** while learning was taking place by defining formative assessment as "the process used by teachers and students to recognise and respond to student learning in order to enhance that learning, **during the learning**" (p. 32, my emphasis). The requirement that the assessment be conducted during learning was also embraced by Shepard, Hammerness, Darling-Hammond, Rust, Snowden, Gordon, Gutierrez and Pacheco (2005) in defining formative assessment as "assessment carried out during the instructional process for the purpose of improving teaching or learning" (p. 275).

The OECD review of formative assessment practices across eight national and provincial systems also emphasised the principle that the assessment should take place during instruction: "Formative assessment refers to frequent, interactive assessments of students' progress and understanding to identify learning needs and adjust teaching appropriately" (Looney, 2005, p. 21). In similar vein, Kahl (2005) wrote: "A formative assessment is a tool that teachers use to measure student grasp of specific topics and skills they are teaching. It's a 'midstream' tool to identify specific student misconceptions and mistakes while the material is being taught" (p. 11).

Broadfoot, Daugherty, Gardner, Gipps, Harlen, James and Stobart (1999) argue that using assessment to improve learning depends on five key factors: 1) the provision of effective feedback to pupils; 2) the active involvement of pupils in their own learning; 3) adjusting teaching to take account of the results of assessment; 4) a recognition of the profound influence assessment has on the motivation and self-esteem of pupils, both of which are crucial influences on learning; and 5) the need for pupils to be able to assess themselves and understand how to improve. They suggest that the term "formative assessment" is unhelpful to describe such uses of assessment because "the

term ‘formative’ itself is open to a variety of interpretations and often means no more than that assessment is carried out frequently and is planned at the same time as teaching” (p. 7). Instead, they suggest the term “assessment for learning”, as proposed originally by James (1992).

Black, Harrison, Lee, Marshall and Wiliam (2004) suggest keeping both terms in that “assessment for learning” refers to any assessment for which the first priority in its design and practice is to serve the purpose of promoting students’ learning, and that this “becomes ‘formative assessment’ when the evidence is actually used to adapt the teaching work to meet learning needs” (p. 10).

Taking this into account, I propose the following definition based on Black and Wiliam (2009), which subsumes and extends previous definitions: “An assessment functions formatively to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of that evidence.”

Several features of this definition are worth noting:

- It is based on the function served by the information yielded by the assessment, rather than a property of the assessment itself.
- The assessment can be carried out by the teacher, the learner, or her peers.
- The focus of the definition is on decisions regarding next steps in instruction, rather than intentions or outcomes.
- The definition is probabilistic.
- The assessment need not change the direction of instruction (it might merely confirm that the planned subsequent actions were appropriate).

Any assessment that provides evidence that has the potential to improve instructional decision-making by teachers, learners, or their peers can therefore be formative. Suppose a class has taken a test that assesses the ability to find the largest or smallest fraction in a given set. The raw scores achieved by students would provide a “monitoring assessment”, indicating which students might benefit from additional instruction or explanation. If, in addition, the teacher noticed that many students gaining low scores were more successful in examples that involved unitary fractions (those with 1 as the numerator) than with more complex fractions, this would provide a “diagnostic assessment”, providing specific information about sources of difficulty. The teacher would then be able to focus additional instruction on non-unitary fractions. If the teacher can see from the responses that many students are operating with

a strategy that the smallest fraction is the one with the largest denominator, and the largest fraction is the one with the smallest denominator – a strategy that works with unitary fractions (Vinner, 1997) – then this provides information for the teacher that is “instructionally tractable”. Such assessments and interpretations of them not only signal the problem (monitoring) and locate it (diagnosing), but they also situate the problem within a theory of action that suggests measures to be taken to improve learning. The best formative assessments are prospective rather than retrospective, therefore, in that they identify recipes for future action.

Any assessment is potentially formative, therefore, since any assessment can support decisions that would not have been possible, or would not be made so well, without the assessment information. However, this does not mean that all formative uses of assessment information are equally effective. By definition, assessments giving diagnostic insights are likely to lead to better decisions about teaching than those that simply monitor student achievement, and those that yield insights that are instructionally tractable are, in all likelihood, better still.

One of the differences between assessments that monitor, those that diagnose, and those that provide insights that are instructionally tractable is the specificity of the information yielded: to be instructionally tractable, the assessment needs to provide more information than simply whether learning is taking place or, if it is not, what specifically is not being learned: it must also incorporate theories of curriculum and of learning. This is because the focus is on “what next?” and this implies a clear notion of a learning progression – a description of the “knowledge, skills, understandings, attitudes or values that students develop in an area of learning, in the order in which they typically develop them” (Forster and Masters, 2004, p.65). Instructional tractability also entails a theory of learning because, before a decision can be made about what evidence to elicit, it is necessary to know not just what comes next in learning, but also what kinds of difficulties learners have in making those next steps. The links between formative assessment and theories of learning are examined in greater detail in Black and Wiliam (2005), Brookhart (2007), Wiliam (2007), and Black and Wiliam (2009).

### *Cycle lengths for formative assessment*

In the example of the fractions test discussed above, the action taken by the teacher follows quickly from generating the evidence about student achievement. In general, however, the definition of formative assessment proposed above allows for cycles of elicitation, interpretation and action of any length, provided the information is used to inform decisions about teaching, which decisions are likely to be better than those made in the absence of that evidence. The length of the formative assessment cycle should also be attuned

to the capacity of the system to respond to the evidence generated – there is little point in generating information on a daily basis if the decisions that the evidence is to inform are taken only monthly (Wiliam and Thompson, 2007).

Not all examples consistent with this definition would be considered as formative assessment under some of the other definitions discussed above. For example, Cowie and Bell (1999), Looney (2005), Shepard (2007) and Kahl (2005) would all probably resist using the term “formative” for assessment that seems remote from its collection. The research literature reviewed above indeed confirms that formative assessment that is less remote is more likely to increase learning and by a greater amount. However, as I have elsewhere noted (Wiliam, 2009), it seems odd to reserve the term “formative” only for assessments that make a significant difference to student outcomes. Rather, it makes more sense to this author to describe assessment as “formative” when it **forms** the direction of future learning but to acknowledge that there are different cycle-lengths involved, as shown in Table 6.2.

Table 6.2. **Cycle lengths for formative assessment**

Type	Focus	Length
Long-cycle	Across marking periods, quarters, semesters, years	4 weeks to 1 year
Medium-cycle	Within and between instructional units	1 to 4 weeks
Short-cycle	Within and between lessons	Day by day: 24 to 48 hours Minute by minute: 5 seconds to 2 hours

Source: Wiliam and Thompson (2007).

### **Formative assessment: key instructional processes**

In order to understand what kinds of formative assessments are likely to be most effective, it is necessary to go beyond the functional definition of formative assessment and look in more detail at the underlying processes. The “systems” metaphor adopted by Ramaprasad (1983), which provides the basis for the definition of assessment for learning adopted by the Assessment Reform Group (Broadfoot *et al.*, 2002), draws attention to three key instructional processes in terms of establishing:

1. Where the learners are in their learning.
2. Where they are going.
3. What needs to be done to get them there.



While many approaches to formative assessment emphasise the role of the teacher, the definition adopted here acknowledges the roles that the learners themselves and their peers have to play. Crossing the process dimension (where learners are in their learning, where they are going, how to get there) with that of the agent in the instructional process (teacher, peer, learner) produces a matrix of nine cells. However, while some of the nine cells generated in this way make sense on their own, it also makes sense to look at other cells in combination. For example, if we consider the role of students in establishing where they are in their learning, and how to reach their desired goal, this can be presented as a process of “activating students as owners of their own learning”, which subsumes a range of important aspects of learning, such as meta-cognition (see Schneider and Stern, this volume). In the same way, the role of peers in establishing where students are in their learning and how they can reach their desired goal, can be presented as “activating students as instructional resources for one another” (see Barron and Darling-Hammond, this volume). Finally, the three cells involving “where the learner is going” can be presented as “clarifying, sharing, and understanding learning intentions and criteria for success”. The result is that the nine cells can be collapsed into the five “classroom strategies” of formative assessment marked 1-5 in Table 6.3. Details of the research base for each of these five strategies can be found in Wiliam (2007), and details of how teachers have implemented these strategies in their own classrooms can be found in Leahy, Lyon, Thompson and Wiliam (2005).

Table 6.3. **Classroom strategies for formative assessment**

	<i>Where the learner is going</i>	<i>Where the learner is right now</i>	<i>How to get there</i>
<i>Teacher</i>	Clarifying learning intentions and sharing and criteria for success (1)	Engineering effective classroom discussions, activities and tasks that elicit evidence of learning (2)	Providing feedback that moves learners forward (3)
<i>Peer</i>	Understanding and sharing learning intentions and criteria for success (1)	Activating students as instructional resources for one another (4)	
<i>Learner</i>	Understanding learning intentions and criteria for success (1)	Activating students as the owners of their own learning (5)	

Source: Leahy, Lyon, Thompson and Wiliam, 2005.

## Formative assessment and the regulation of learning processes

In the remainder of this chapter, I discuss how the approach to formative assessment outlined here can be integrated into a larger perspective on instructional design through a focus on the regulation of learning processes (Perrenoud, 1991; 1998).

Within such a framework, the actions of the teacher, the learners, and the context of the classroom can be evaluated with respect to how well the intended learning proceeds towards the intended goal. As Schneider and Stern (this volume) point out, teachers do not create learning; only learners can do this and so many have called for a shift in the role of the teacher from the “sage on the stage” to the “guide on the side.” The danger with such a characterisation is that it is often interpreted as relieving the teacher of responsibility for ensuring that learning takes place. What I propose here is that the teacher be regarded as responsible for “engineering” a learning environment, both in its design and its operation.

An effective learning environment creates student engagement and is well-regulated. As a growing body of research on cognitive development shows, the level of engagement in cognitively challenging environments influences not only achievement, but also IQ itself (Dickens and Flynn, 2001; Mercer, Dawes, Wegerif and Sams, 2004). As well as creating engagement, effective learning environments need to be designed so that, as far as possible, they afford or scaffold the learning that is intended (“proactive regulation”). If the intended learning is not occurring, then this should become apparent so that appropriate adjustments may be made (“interactive regulation”). Finally, it is also possible for teachers to engage in “retroactive regulation”; for example, when a teacher realises that a particular instructional sequence might be improved for one group of students as a result of experiences with other groups of students.

Proactive regulation is achieved “upstream” of the lesson itself (*i.e.* before the lesson begins). The regulation can be unmediated as when, for example, a teacher “does not intervene in person, but puts in place a ‘meta-cognitive culture’, mutual forms of teaching and the organisation of regulation of learning processes run by technologies or incorporated into classroom organisation and management” (Perrenoud, 1998, p. 100). For example, a teacher’s decision to use realistic contexts in mathematics can provide a source of regulation since students will be able to evaluate how reasonable are their answers. When a teacher develops in the students the skills of consulting and productively supporting each other, this too is an example of proactive regulation.

At other times, particularly when it is hard to predict how students will respond to instructional activities, it may be more appropriate to regulate learning interactively – for example, by creating questions, prompts or

activities that evoke responses from the students that the teacher can use to determine the progress of the learning and, if necessary, to make adjustments. Often, these questions or prompts will be open-ended, requiring higher-order thinking – indeed such questions are essential to creating learning environments that foster student engagement. But closed questions have a role here, too. “Is calculus exact or approximate?”, “What is the pH of 10 molar NaOH?”, or, “Would your mass be the same on the moon?” are all closed questions with a single correct answer, but are valuable because they frequently reveal student conceptions that are different from those intended by the teacher (many students believe that calculus is approximate, that a pH cannot be greater than 14, and that one’s mass depends on gravity like one’s weight does).

“Upstream” planning of good questions like those above therefore creates the possibility that the learning activities “downstream” may change course in light of the students’ responses. These “moments of contingency” – points in the instructional sequence when the instruction can proceed in different directions according to the responses of the students – are at the heart of the regulation of learning. Indeed, Black and Wiliam (2009) propose that formative assessment is, in essence, concerned with “the creation of, and capitalisation upon, ‘moments of contingency’ in instruction for the purpose of the regulation of learning processes” (p. 6). A theory of formative assessment is therefore much narrower than an overall theory of teaching and learning, although it links in significant ways to other aspects of teaching and learning, since how teachers, learners, and their peers create and capitalise on these moments of contingency entails considerations of instructional design, curriculum, pedagogy, psychology and epistemology.

## Summary

This chapter has traced a number of significant strands in the development of the concept of formative assessment, although the account is of necessity highly selective. The earliest uses of the term drew heavily on the idea of feedback and on navigational metaphors, focusing on feedback as a corrective measure to restore learning to its intended trajectory. Over the last hundred years, literally thousands of studies have sought to determine what kinds of feedback interventions improve learning, and by how much, but these studies are of limited value due to weak conceptualisation of the feedback intervention itself, of the kinds of learning under study, and a failure to consider long-term impacts. Over the last twenty years, there has been considerable interest in the use of formative assessment not in isolation but as an integral feature of high-quality educational practice in classroom settings, and a number of definitions have been proposed.

In this chapter, a definition of formative assessment has been presented emphasising the role of assessment in improving the quality of instructional decisions, which subsumes previous definitions of “formative assessment”. Consequences of this definition have been drawn out; specifically, it is suggested that formative assessment can usefully be thought of as entailing five key strategies:

1. Clarifying, sharing and understanding learning intentions and criteria for success.
2. Engineering effective classroom discussions, activities and tasks that elicit evidence of learning.
3. Providing feedback that moves learners forward.
4. Activating students as instructional resources for one another.
5. Activating students as the owners of their own learning.

Finally, it is suggested that formative assessment is concerned with the creation of, and capitalisation upon, “moments of contingency” in instruction with a view to regulating learning processes, which allows a clear demarcation between formative assessment and other aspects of instructional design and pedagogy.

## Notes

1. They began by identifying approximately 3 000 potentially relevant research studies, and excluded all those with fewer than 10 participants, where there was not a comparison group of some kind, and those with too few details for effect sizes to be computed. They were left with just 131 publications, reporting 607 effect sizes and involving 23 663 observations of 12 652 participants.
2. From an initial screening involving on-line databases which generated 180 relevant studies, a total of 141 publications met the inclusion criteria (103 journal articles, 24 books and book chapters, 10 conference proceedings and 4 research reports).
3. The French word *régulation* has a much more specific meaning than the English word “regulation”. There are two ways to translate the word “regulation” into French – *règlement* and *régulation*. The former of these is used in the sense of “rules and regulations,” while the latter is used in the sense of adjustment in the way that a thermostat “regulates” the temperature of a room.

## References

- Allal, L. and L.M. Lopez (2005), “Formative Assessment of Learning: A Review of Publications in French” in J. Looney (ed.), *Formative Assessment: Improving Learning in Secondary Classrooms*, Paris, France: Organisation for Economic Cooperation and Development, pp. 241-264.
- Ausubel, D.P. (1968), *Educational Psychology: A Cognitive View*, Holt, Rinehart and Winston, New York.
- Bangert-Drowns, R., C. Kulik, J. Kulik and M. Morgan (1991), “The Instructional Effect of Feedback in Test-Like Events”, *Review of Educational Research*, Vol. 61, No. 2, pp. 213-238.
- Black, P., C. Harrison, C. Lee, B. Marshall and D. Wiliam (2004), “Working Inside the Black Box: Assessment for Learning in the Classroom”, *Phi Delta Kappan*, Vol. 86, No. 1, pp. 9-21.
- Black, P. and D. Wiliam (1998a), “Assessment and Classroom Learning”, *Assessment in Education: Principles Policy and Practice*, Vol. 5, No. 1, pp. 7-73.
- Black, P. and D. Wiliam (2005), “Developing a Theory of Formative Assessment”, in J. Gardner (ed.), *Assessment and learning*, Sage, London, UK, pp. 81-100.
- Black, P. and D. Wiliam (2009), “Developing the Theory of Formative Assessment”, *Educational Assessment, Evaluation and Accountability*, Vol. 21, No. 1, pp. 5-31.
- Bloom, B.S. (1969), “Some Theoretical Issues Relating to Educational Evaluation”, in R.W. Tyler (ed.), *Educational Evaluation: New Roles, New Means: The 68th Yearbook of the National Society for the Study of Education (part II)* (Vol. 68, No. 2, pp. 26-50), University of Chicago Press, Chicago, IL.
- Broadfoot, P., R. Daugherty, J. Gardner, W. Harlen, M. James and G. Stobart (1999), *Assessment for Learning: Beyond the Black Box*, University of Cambridge School of Education, Cambridge, UK.

- Broadfoot, P., R. Daugherty, J. Gardner, W. Harlen, M. James and G. Stobart (2002), *Assessment for Learning: 10 principles*, University of Cambridge School of Education, Cambridge, UK.
- Brookhart, S.M. (2004), “Classroom Assessment: Tensions and Intersections in Theory and Practice”, *Teachers College Record*, Vol. 106, No. 3, pp. 429-458.
- Brookhart, S.M. (2007), “Expanding Views about Formative Classroom Assessment: A Review of the Literature”, in J.H. McMillan (ed.), *Formative Classroom Assessment: Theory into Practice*, Teachers College Press, New York, NY, pp. 43-62.
- Cowie, B. and B. Bell (1999), “A Model of Formative Assessment in Science Education”, *Assessment in Education: Principles Policy and Practice*, Vol. 6, No. 1, pp. 32-42.
- Crooks, T.J. (1988), “The Impact of Classroom Evaluation Practices on Students”, *Review of Educational Research*, Vol. 58, No. 4, pp. 438-481.
- Deci, E.L. and R.M. Ryan (1994), “Promoting Self-Determined Education”, *Scandinavian Journal of Educational Research*, Vol. 38, No. 1, pp. 3-14.
- Dempster, F.N. (1991), “Synthesis of Research on Reviews and Tests”, *Educational Leadership*, Vol. 48, No. 7, pp. 71-76.
- Dempster, F.N. (1992), “Using Tests to Promote Learning: A Neglected Classroom Resource”, *Journal of Research and Development in Education*, Vol. 25, No. 4, pp. 213-217.
- Denvir, B. and M.L. Brown (1986a), “Understanding of Number Concepts in Low-Attaining 7-9 year olds: Part 1. Development of Descriptive Framework and Diagnostic Instrument”, *Educational Studies in Mathematics*, Vol. 17, No. 1, pp. 15-36.
- Denvir, B. and M.L. Brown (1986b), “Understanding of Number Concepts in Low-Attaining 7-9 year olds: Part II, The Teaching Studies”, *Educational Studies in Mathematics*, Vol. 17, No. 2, pp. 143-164.
- Dickens, W. and J.R. Flynn (2001), “Heritability Estimates vs. Large Environmental Effects: The IQ Paradox Resolved”, *Psychological Review*, Vol. 108, No. 2, pp. 346-369.
- Elshout-Mohr, M. (1994), “Feedback in Self-Instruction”, *European Education*, Vol. 26, No. 2, pp. 58-73.
- Forster, M. and G. Masters (2004), “Bridging the Conceptual Gap between Classroom Assessment and System Accountability”, in M. Wilson (ed.), *Towards Coherence between Classroom Assessment and Accountability*, University of Chicago Press, Chicago.

- Fuchs, L.S. and D. Fuchs (1986), “Effects of Systematic Formative Evaluation – A Meta-Analysis”, *Exceptional children*, Vol. 53, No. 3, pp. 199-208.
- Harlow, L.L., S.A. Mulaik and J.H. Steiger (eds.) (1997), *What If There Were No Significance Tests?*, Lawrence Erlbaum Associates, Mahwah, NJ.
- Hattie, J. and H. Timperley (2007), “The Power of feedback”, *Review of Educational Research*, Vol. 77, No. 1, pp. 81-112.
- James, M. (1992), “Assessment for Learning”, annual conference of the Association for Supervision and Curriculum Development (Assembly session on ‘Critique of Reforms in Assessment and Testing in Britain’) held in New Orleans, LA, University of Cambridge Institute of Education, Cambridge, UK.
- Kahl, S. (2005), “Where in the World are Formative Tests? Right under Your Nose!”, *Education Week*, Vol. 25, No. 4, pp. 11.
- Kluger, A.N. and A. DeNisi (1996), “The Effects of Feedback Interventions on Performance: A Historical Review, A Meta-Analysis, and A Preliminary Feedback Intervention Theory”, *Psychological Bulletin*, Vol. 119, No. 2, pp. 254-284.
- Köller, O. (2005), “Formative Assessment in Classrooms: A Review of the Empirical German Literature”, in J. Looney (ed.), *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD Publishing, Paris, pp. 265-279.
- Leahy, S., C. Lyon, M. Thompson and D. Wiliam (2005), “Classroom Assessment: Minute-by-Minute and Day-by-Day”, *Educational Leadership*, Vol. 63, No. 3, pp. 18-24.
- Looney, J. (ed.) (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD Publishing, Paris.
- Meisels, S.J., S. Atkins-Burnett, Y. Xue, J. Nicholson, D.D. Bickel and S.-H. Son (2003), “Creating a System of Accountability: The impact of Instructional Assessment on Elementary Children’s Achievement Test Scores”, *Education Policy Analysis Archives*, Vol. 11, No. 9.
- Mercer, N., L. Dawes, R. Wegerif and C. Sams (2004), “Reasoning as a Scientist: Ways of Helping Children to Use Language to Learn Science”, *British Educational Research Journal*, Vol. 30, No. 3, pp. 359-377.
- Natriello, G. (1987), “The Impact of Evaluation Processes on Students”, *Educational Psychologist*, Vol. 22, No. 2, pp. 155-175.
- Nyquist, J.B. (2003), *The Benefits of Reconstructing Feedback as a Larger System of Formative Assessment: A Meta-Analysis*, Unpublished Master of Science, Vanderbilt University.

- Perrenoud, P. (1991), “Towards a Pragmatic Approach to Formative Evaluation”, in P. Weston (ed.), *Assessment of Pupils’ Achievement: Motivation and School Success*, Amsterdam: Swets & Zeitlinger, pp. 77-101.
- Perrenoud, P. (1998), “From Formative Evaluation to a Controlled Regulation of Learning Towards a Wider Conceptual Field”, *Assessment in Education: Principles Policy and Practice*, Vol. 5, No. 1, pp. 85-102.
- Ramaprasad, A. (1983), “On the Definition of Feedback”, *Behavioural Science*, Vol. 28, No. 1, pp. 4-13.
- Sadler, D.R. (1989), “Formative Assessment and the Design of Instructional Systems”, *Instructional Science*, Vol. 18, No. 2, pp. 119-144.
- Shepard, L.A. (2007), “Formative Assessment: Caveat Emptor”, in C.A. Dwyer (ed.), *The Future of Assessment: Shaping Teaching and Learning*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 279-303.
- Shepard, L., K. Hammerness, L. Darling-Hammond and F. Rust (2005), “Assessment”, in L. Darling-Hammond and J. Bransford (eds.), *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*, Jossey-Bass, San Francisco, CA, pp. 275-326.
- Shute, V.J. (2008), “Focus on Formative Feedback”, *Review of Educational Research*, Vol. 78, No. 1, pp. 153-189.
- Torrance, H. and J. Pryor (1998), *Investigating Formative Assessment*, Open University Press, Buckingham, UK.
- Tunstall, P. and C.V. Gipps (1996a), “Teacher Feedback to Young Children in Formative Assessment: A Typology”, *British Educational Research Journal*, Vol. 22, No. 4, pp. 389-404.
- Tunstall, P. and C.V. Gipps (1996b), “‘How Does Your Teacher Help You to Make Your Work Better?’ Children’s Understanding of Formative Assessment”, *The Curriculum Journal*, Vol. 7, No. 2, pp. 185-203.
- Vinner, S. (1997), From Intuition to Inhibition – Mathematics, Education and Other Endangered Species, in E. Pehkonen (ed.), *Proceedings of the 21<sup>st</sup> Conference of the International Group for the Psychology of Mathematics Education* (Vol. 1), University of Helsinki Lahti Research and Training Centre, Lahti, Finland, pp. 63-78.
- Wiener, N. (1948), *Cybernetics, or the Control and Communication in the Animal and the Machine*, John Wiley, New York, NY.
- William, D. (2007), “Keeping Learning on Track: Classroom Assessment and the Regulation of Learning”, in F.K. Lester Jr (ed.), *Second Handbook*

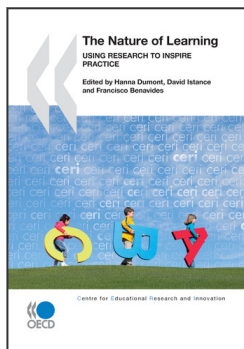


*of Mathematics Teaching and Learning*, Information Age Publishing, Greenwich, CT, pp. 1053-1098.

Wiliam, D. (2009), "An Integrative Summary of the Research Literature and Implications for a New Theory of Formative Assessment", in H.L. Andrade and G.J. Cizek (eds.), *Handbook of Formative Assessment*, Routledge, Taylor and Francis, New York.

Wiliam, D. and P.J. Black (1996), "Meanings and Consequences: A Basis for Distinguishing Formative and Summative Functions of Assessment?", *British Educational Research Journal*, Vol. 22, No. 5, pp. 537-548.

Wiliam, D. and M. Thompson (2007), "Integrating Assessment with Instruction: What Will it Take to Make it Work?", in C.A. Dwyer (ed.), *The Future of Assessment: Shaping Teaching and Learning*, Lawrence Erlbaum Associates, Mahwah, NJ, pp. 53-82.



**From:**  
**The Nature of Learning**  
Using Research to Inspire Practice

**Access the complete publication at:**  
<https://doi.org/10.1787/9789264086487-en>

**Please cite this chapter as:**

William, Dylan (2010), “The role of formative assessment in effective learning environments”, in Hanna Dumont, David Istance and Francisco Benavides (eds.), *The Nature of Learning: Using Research to Inspire Practice*, OECD Publishing, Paris.

DOI: <https://doi.org/10.1787/9789264086487-8-en>

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to [rights@oecd.org](mailto:rights@oecd.org). Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at [info@copyright.com](mailto:info@copyright.com) or the Centre français d'exploitation du droit de copie (CFC) at [contact@cfcopies.com](mailto:contact@cfcopies.com).