

18

DYLAN WILIAM

ASSESSMENT FOR LEARNING

LEARNING OUTCOMES

Having read this chapter you should be able to:

- appreciate Dylan Wiliam's work on assessment for learning and understand how this can be applied in the classroom
- understand the difference between assessment for learning and assessment of learning, and recognise the implications for pupils
- critically appraise his work.

KEY WORDS

assessment for learning; Black Box; formative assessment; Teacher Learning Communities; summative assessment; assessment of learning

INTRODUCTION

Dylan Wiliam is perhaps most well-known for his pioneering work on **assessment for learning**, work undertaken with his colleague Paul Black, which culminated in the best-selling publication *Inside the Black Box* (Black and Wiliam, 1998a). Currently Emeritus Professor of Educational Assessment at University College London, Wiliam's career began as a secondary school teacher, initially as a tutor of mathematics and physics at a private school in Worcestershire, followed by teaching positions in schools in inner-city London. His career in academia began in 1984 when he took a two-year research fellow post at Chelsea College working on a scheme called Graded Assessment in Mathematics – a precursor to **formative assessment**.

Wiliam's academic career has seen him combine his early experience as a classroom teacher with research work into education systems, initially in assessment practices and latterly in looking at the professional development of teachers, which led to his work around the development of **Teacher Learning Communities** (TLCs) in which he presented a vision for a support group for teachers. Wiliam proposes that TLCs should encourage teachers to choose an area of practice which they wish to change, and colleagues within the group support individuals in making these changes. He explains:

What you as a teacher are trying to learn comes from your own needs – so it's a bit like Alcoholics Anonymous or Weightwatchers – the support group aren't there to tell you what to do, but to help you achieve what you want to do. (Wiliam, cited in Goodwin, 2008: 4)

Wiliam's TLCs were particularly successful in Southern Australia, Singapore and Sweden and he currently divides his time between these countries and the UK.

McInerney (2015) suggests that Wiliam is 'dangerously close to being a public intellectual' and observes that he is 'one of the few modern educationalists to become a familiar staffroom name'. This may be in part down to his appearance on the 2010 BBC documentary *The Classroom Experiment*, in which he introduced his teaching methods to a group of Hertfordshire teachers. In this documentary he attempted to demonstrate how subtle changes to teaching methods could reap rewards, and advised that such methods were more effective than spending vast sums of money on school buildings and reducing class sizes – arguing that instead investment should be in raising teachers' skill levels.

Wiliam's work has shown a commitment to improving aspects of the education system in the best interests of both children and practitioners, and while some of his ideas have proved controversial, it cannot be denied that he has had a significant impact on teaching practices, especially in the area of whole-school assessment and feedback.

DYLAN WILIAM, THE PERSON

Dylan Ap Rhys Wiliam was born in North Wales, the son of an Oxford-educated professor. Wiliam was brought up as a Welsh speaker and did not really speak any English until he went to grammar school in Cardiff and Manchester (Wilby, 2011). He admits to being a poor student, stating 'I was very badly co-ordinated and couldn't write legibly. I still print, basically' (McInerney, 2015). He has also reflected on being badly behaved, uncoordinated and having a stammer, and always being the last to be picked for teams in physical education (Wilby, 2011). Fortunately for Wiliam he found that he had a talent for mathematics in his late teens and was awarded a school maths prize for the best A-level results. In addition, he also began weight training at the age of fourteen, and after training for an hour every day for three years he explained that he was transformed from 'being completely useless at sports to being a kind of jock' (McInerney, 2015), making captain of the rugby team and house captain in athletics.

Wiliam's average education performance continued throughout university, and after failing to get an interview for both Oxford and Cambridge he accepted a place at Durham studying mathematics and physics. Of his time at Durham, he explains, 'I don't know if I was depressed at university or not, but there are large areas where there is a sort of haze about what actually happened' (McInerney, 2015). Wiliam eventually graduated with a third-class honours degree. While at university he became heavily involved with music, and this distraction may well further explain his poor academic performance. He played in a 'jazz-folk' band called Lynx, and wholly intended to pursue a musical career on graduation. However, McInerney (2015) states that his plans to pursue a career in music were impeded after his father took a job at the University of Connecticut and informed Wiliam that he would not be able to remain in the family home.

Wiliam decided to apply for teaching as a means of supporting himself while he continued to play in his band; however, after failing to get a place on a Postgraduate Certificate in Education (PGCE) course at Goldsmiths, he went to an employment agency specialising in recruiting teachers to private schools, where he was offered a job working at a private school in Worcestershire. It was at this time that he realised his musical ambitions would be unfulfilled since occasional gigs in pubs had proved insufficient to make a living. He reflected that 'I hadn't realised how hard it is to make any kind of a living as a musician. If you get enough money to cover your petrol, you're lucky' (Wilby, 2011). Fortunately for Wiliam this realisation came around the same time that he determined that he actually enjoyed teaching, and thus he decided to dedicate himself to this profession. In addition, he decided to move out of the private and into the state sector, working in secondary schools in West London while at the same time gaining further degrees in mathematics and mathematics education through the Open University and the Polytechnic of the South Bank.

William taught in London secondary schools between 1976 and 1984 and quickly gained promotions, in part because he was teaching a 'shortage subject'. However, after eight years of teaching, and following a disagreement with his head teacher over a promotion opportunity, he joined Chelsea College to work on a research project involving Graded Assessment in Mathematics. He has reflected that at the time everyone thought he was being somewhat rash, leaving a secure job for a two-year fixed contract; however, he has responded to this by saying that he jumped at the chance to work on something that interested him in greater depth (Goodwin, 2008). While William suggests that research was not something he wanted to do, his time at Chelsea College, later to become a part of King's College in London, led to further roles at King's as a lecturer in mathematics, Dean of the School of Education and then Assistant Principal. It was during this period that he studied for his doctorate, which he gained in 1993 from the University of London.

It was during his time at King's College that William first worked with Paul Black, a relationship which he describes as 'intellectually the most productive of my career' (Goodwin, 2008). William believes that the success of their working relationship is due to the fact that they work very differently and, having different areas of expertise (Black is a physicist), they have been forced to think about things, and make connections that they would not have done had they worked independently. It was his work on formative assessment with Paul Black that led to his most well-known work, *Inside the Black Box*, and while William jokes that this made them what appeared to be an overnight success, the work actually took over twenty years to come to fruition.

It is possibly due to his early career as a teacher that William has dedicated his academic career to improving classroom practice and supporting classroom teachers. He believes that one of the reasons for the success of *Inside the Black Box* is the embedding of formative assessment into practice rather than just through policy (Goodwin, 2008). He has spoken out against a results-driven agenda and has sought ways of supporting high-quality teaching through methods which will be further explored in the application section of this chapter. Despite not having a traditional teaching degree he appears to have gained a 'guru-like' status among the teaching profession; however, he remains humble, and reflecting on his academic career he says 'certainly for a long time in my career I had this imposter syndrome thing of "they will find out there has been a mistake and I will have to go back to teaching in Shepherd's Bush"' (McInerney, 2015). However, he also shows the courage of his convictions, refusing to do something just because others expect him to. This can be seen reflected in his work with teachers, which frequently shows him throwing out the rulebook – as we will see later in the chapter.

THE BLACK BOX

As we have seen earlier in the chapter, William's most influential work was a result of studies into the impact of assessment on pupils, culminating in the publication *Inside*

the Black Box (Black and Wiliam, 1998a). For Black and Wiliam, the Black Box referred to the classroom in which outside inputs from policy makers are fed in with the hope that this will result in desirable outputs. For example, policy related to testing and assessment should result in better outcomes for pupils and teachers. However, Black and Wiliam raised the question, 'how can anyone be sure that a particular set of new inputs will produce better outputs if we don't at least study what happens inside?' (1998a: 1). Black and Wiliam went on to suggest that the impetus was on teachers to make the inside work better. However, they saw this as an unfair system, especially in areas as important as standards raising. They suggested that policy makers should take on more responsibility for providing direct help – setting out their recommendations in the aforementioned publication.

Black and Wiliam had a specific interest in assessment practices, defining assessment as 'all those activities undertaken by teachers, and by their students in assessing themselves, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged' (Black and Wiliam, 1998a: 2). For them, it was not just the end result that was important, but also the learning process; key to the success of assessment were the ongoing changes which were made as a result of the observations made. Black and Wiliam called this process formative assessment in which 'the evidence is actually used to adapt the teaching work to meet the needs' (1998a: 2). In forming their conclusions, they conducted a longitudinal project, reviewing research in books and journal articles over a nine-year period. Their results were first published in the journal *Assessment in Education* (1998b), in which they drew the conclusion that formative assessment did indeed raise standards. However, there was room for improvement, especially in relation to formative assessment practices.

Formative assessment is a complex process, far more so than **summative assessment**, which Black and Wiliam suggested formed the basis of public and political attention. At the time of the Black Box publication, assessment had been very much concentrated around the tests conducted at the end of each key stage, with a focus on the overall levels and grades. Black and Wiliam acknowledged that teachers did make some contribution to these, but also suggested that little attention was paid to this. Summative assessment then would be interpreted as an **assessment of learning**, referring to an assessment type which yields a final result – for example, numerical data to be published in league tables, or certificates given to students such as GCSEs which give employers or education providers an indication of a pupil's performance. While valid in terms of providing a reliable indication of pupil performance, comparable across the country, this type of assessment forms no useful purpose in promoting pupil learning or in informing teachers of how they might move that learning forward. Summative assessment normally takes place at the end of a learning unit or module, by which time it is often too late to address any difficulties pupils may have experienced.

Formative assessment, on the other hand, is far more effective as a diagnostic tool, allowing for difficulties to be seen and addressed prior to summative assessments

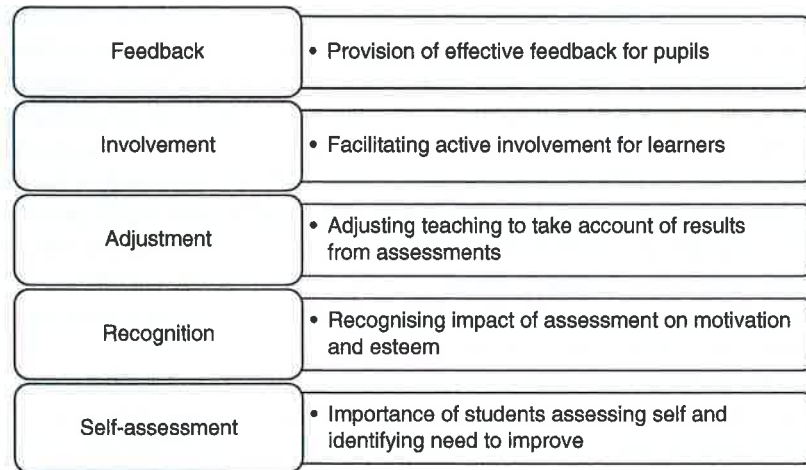


Figure 18.1 Elements identified for assessment to improve learning

being carried out. Following on from *Inside the Black Box*, the Assessment Reform Group (ARG) (1999), of which Dylan Wiliam was a part, set out five elements for assessment to improve learning in the follow-up publication *Assessment for Learning: Beyond the Black Box* (see Figure 18.1).

It was also at this point that the term ‘formative assessment’ was rejected in favour of *assessment for learning*, since it was considered that formative assessment was open to interpretation and simply suggested that assessment was carried out and planned at the same time as teaching, negating the notion that it should be used to inform planning. Nevertheless, the fundamental principles behind formative assessment remained, and the five elements as set out by the ARG (see Figure 18.1) will now be considered in more detail.

PRINCIPLE 1. EFFECTIVE FEEDBACK

Black and Wiliam (1998a) observed that teacher feedback to pupils frequently served a social and managerial function with little attention paid to the learning function it should play. Wiliam (2016) observes that the traditional approach to feedback came through the marking of work, normally after the event, and follows three steps: firstly the teacher looks at the pupil’s work, this is then followed by feedback of some sort, either a grade or mark and sometimes a comment, and finally the pupil looks at the feedback, usually after the event. While ostensibly the purpose of feedback is to improve pupil learning, Wiliam suggested that ‘much of the feedback that students get has little or no effect on their learning, and some kinds of feedback are actually counterproductive’ (2011: 107). Furthermore, as observed by Black et al. (2003), this type

of marking rarely offered advice as to how the work could be improved and often reinforced underachievement. Moreover, pupils were rarely given sufficient time to address any shortcomings or misconceptions.

For Black and Wiliam quality feedback 'should be about the particular qualities of his or her work, with advice on what he or she can do to improve and should avoid comparisons with other pupils' (1998a: 9). Wiliam (2016) proposes that feedback should indicate the gap between current performance and desired performance and should improve a pupil's performance in tasks they have not yet attempted. That is, feedback should not simply tell the pupil what they need to do to improve a specific piece of work, but should enable them to develop the skills necessary to improve on subsequent pieces of work. Furthermore, the most effective feedback is that which takes place during the event, giving pupils the opportunity to address it immediately, and where this is not possible feedback should be given soon after the event and time should be given for pupils to read and address the feedback given, since as Wiliam observes, 'the only important thing about feedback is what students do with it' (2016: 10).

PRINCIPLE 2. ACTIVE INVOLVEMENT

Black and Wiliam (1998a) advised that pupils should be involved in the choice of tasks, inasmuch as prior learning is considered at the planning process. Additionally, the sharing of learning aims is encouraged, with pupils having the opportunity to 'communicate their evolving understanding' (1998a: 10). Black and Wiliam (1998a) suggested that time should be set aside for this, also encouraging pupil-teacher dialogue, which might allow teachers to respond to and reorientate pupils' thinking without inhibiting future learning. Black and Wiliam recommended that 'the dialogue between pupils and a teacher should be thoughtful, reflective, focused to evoke and explore understanding, and conducted so that all pupils have an opportunity to think and to express their ideas' (1998a: 12).

PRINCIPLE 3. ADJUSTMENT TO TEACHING

Black and Wiliam (1998a) did not reject the notion of summative assessment; however, for them the end product of the assessment – the result – was less important than what a teacher could learn from the result: hence the third principle promotes the adjustment of teaching to take account of the result of the assessment. In this instance, summative assessment is used in a formative way. In *Inside the Black Box* (1998a) they outlined that a good test should be a learning as well as a testing opportunity. The ARG (1999) advised that high-stakes testing, such as end of key stage tests, had resulted in teachers adopting practice testing rather than focusing on assessment for learning, which was resulting in pupil anxiety, particularly among lower achievers. An

alternative approach, advocated by Black et al. (2003), was to 'use the aftermath of tests as an opportunity for formative work' (2003: 55). One such approach might be for teachers to analyse test papers and identify which questions were completed incorrectly by a majority of pupils; attention could subsequently be concentrated on these concepts in future planning, while problems encountered by a smaller minority could be addressed through individual or small-group tutoring.

PRINCIPLE 4. SELF-ESTEEM AND MOTIVATION

The fourth principle is related to the influence of self-esteem and motivation of pupils, with Black et al. (2003) observing that an overemphasis on testing and grades can lead to comparisons between students, resulting in the emergence of a competitive element to classroom practice as opposed to the self-improvement which should be advocated. In this principle a correlation can be seen with the use of feedback, with Black et al. (2003) suggesting that feedback involving grades and scores enhances ego rather than promoting task engagement. They go on to suggest that pupils focus on image and status, rather than reflections on the work itself and how they can improve it. Moreover, for low attainers, a focus on grades and scores can be damaging to their self-esteem. It is suggested therefore that feedback which focuses on next steps and 'what needs to be done' sends a clear message that improvements can be made, increasing self-esteem and improving motivation as the pupil increases their efforts to make the necessary changes (Black et al., 2003: 46). Black et al. (2003) propose that this promotes a culture of success in which all pupils, regardless of ability, can see how achievements can be made through building on their own previous performance, as opposed to competing with their peers.

PRINCIPLE 5. SELF-ASSESSMENT

In the final element the ARG outlined the importance of pupils being involved in self-assessment and understanding how they can use this to improve their own learning. Wiliam (2011) expressed this as students being owners of their own learning, reinforcing the notion that it is the learners and not the teachers who should create the learning experience. However, Wiliam also reflected that the school culture appeared to 'be based on the opposite principle that if teachers try really hard, they can do the learning for the learners' (2011: 169). In this scenario the pupils become passive recipients of their own learning. From the research which informed *Inside the Black Box*, Wiliam made the point that, if allowed, pupils are more than capable of developing sufficient insights into their own learning to make the necessary improvements, and furthermore, evidence suggests that such involvement can also improve pupils' motivation. Black and Wiliam (1998a) advised that pupils should be trained in self-assessment, thereby giving them ownership of their own learning and recognising for

themselves what they need to do to achieve. Black and Wiliam (1998a) also suggested that pupils were very capable of assessing their own work and tended to be both honest and reliable in their assessments of their own performance and that of their peers – in fact they were frequently much harder on themselves than their teachers. However, Black and Wiliam (1998a) also proposed that self-assessment could only happen if the pupils were aware of what they were trying to achieve, hence the importance of sharing targets and learning outcomes with pupils. There was no misapprehension of the challenges of this aspect of formative assessment, since a tradition of passive receptivity had been ingrained in the education system. However, Black and Wiliam observed that self-assessment by pupils, ‘far from being a luxury, is in fact an essential component of formative assessment’ (1998a: 10) and well worth the effort of teaching pupils how to carry this out effectively.

It should be noted that the application of formative assessment to a school system which had come to rely on more traditional methods was always going to be a challenge. However, Black and Wiliam (1998a) determined that there was a need for an increased focus on the inside of the Black Box which could best be achieved through a change to policy. Some evidence of the enduring status of the work of the ARG can be seen in a 2018 Department for Education (DfE) review undertaken by the Commission on Assessment Without Levels (CAWL), an advisory group established following the removal of National Curriculum levels in 2014. In their review of assessment practices (DfE, 2018) CAWL reported that formative assessments were commonly used by schools to ‘evaluate pupils’ knowledge and understanding on a day-to-day basis and to tailor teaching accordingly’ (2018: 20), with the most commonly used methods including ‘quizzes, “low-stakes” skills/knowledge tests, “thumbs up or thumbs down”, questioning in class, pupil self-assessment, homework, marking and “open-book” tasks’ (2018: 21). As we will see in the application to practice section of this chapter, many of the methods seen here formed the basis of the practical advice for classroom practitioners which became a key focus for later work undertaken by Wiliam.

LINKS WITH OTHER RESEARCHERS

In formulating the original *Inside the Black Box* material Black and Wiliam (1998a) conducted an extended literature review of research into assessment, commencing with 580 articles or chapters of which 250 sources were eventually chosen to be examined in more detail. In view of this it can be seen that their work was influenced by a significant number of researchers in the field, some of whom will be explored in this section.

Wiliam (2014a) proposed that Michael Scriven was the first to use the term ‘formative assessment’ as early as 1967, but as a means of ongoing improvement of the curriculum as opposed to the assessment of individual pupils. This was further developed by Benjamin Bloom (1969), predominantly in the higher education context,

where formal assessments or tests were used to inform future teaching or curriculum management. However, Wiliam (2014a) suggests that Sadler (1989) was the most influential in terms of seeing 'formative assessment as being intrinsic to, and integrated with, effective instruction' (2014a: 2). Sadler defines formative assessment as being concerned with how judgements about the quality of students' responses (performances, pieces of work) can be used to shape or improve their competence by short-circuiting the randomness and inefficiency of trial and error learning (1989: 120). Black and Wiliam (1998a) used Sadler's observations on the role of the student in their own learning to develop their ideas on peer and self-assessment. Sadler (1989) argued that students could only achieve a learning goal if they could understand that goal and recognise how to achieve it – a concept which is inherent in the Assessment for Learning (AfL) guidance produced by the Department for Children, Schools and Families (DCSF, 2009) in response to Black and Wiliam's work, which sees teachers sharing learning goals with pupils, and offering guidance as to what a successful outcome might look like.

Ruth Butler's (1988) work on feedback to pupils influenced Black and Wiliam's work on closing the gap feedback. Butler (1988) observed that pupils tended to ignore feedback when marks were given, resulting in their being mark driven and not seeking to find ways of improving that work through engagement with the written feedback. Black and Wiliam further developed these observations, and through working with pupils and teachers they proposed that feedback and marks should not be used together, and that if marks were to be given this should happen after pupils had had sufficient time to read and address the feedback given.

Finally, Black and Wiliam (1998a) were influenced by Rowe's (1974) work on the use of questioning in the classroom. Rowe identified that teachers did not give pupils sufficient time to respond to teachers' questions (usually approximately 0.9 seconds). Conducting her own research, Rowe (1974) found that increasing the 'waiting time' had positive effects on pupil engagement, including longer and more confident responses, a decrease in failure to respond and an increased chance of pupils challenging or improving on their own answers. Black and Wiliam's own research led to further suggestions around the use of questioning to include the increasing of wait time, but also in the framing of questions to allow pupils opportunities to explore their own understanding of a concept (Black et al., 2003).

CRITIQUING WILIAM

Wiliam is one of the first to admit that the response of the government to the *Inside the Black Box* research left a lot to be desired (Goodwin, 2008), and suggests that the challenge of encouraging people to change an already entrenched way of thinking was one of the reasons why he and Paul Black were not as successful in promoting their ideas as they might have been. Likewise, Wiliam acknowledges that the exposure

he received as a result of the BBC2 documentary could have come across as gimmicky, agreeing that the 'programme failed to convey a coherent message' (Wilby, 2011). These views are supported by education blogger Joe Kirby (2013), who suggests that formative assessment is a distorted educational concept, backed by a range of gimmicks, unhelpful acronyms and government-led policy which confused AfL with national levels. Kirby also posits that school leadership teams who stuck rigidly to the 'letter of the AfL law' (2013) failed to embrace the flexibility of formative assessment, and in so doing allowed the spirit of what Black and Wiliam (1998a) were trying to achieve to be lost. Indeed, the very fact that Ofsted began to focus part of their observations on establishing how far pupils understood what levels they were at, suggests a system which was far removed from what Black and Wiliam originally intended.

Wiliam refers to this as *policy diffraction* (Menzies, 2010) and uses this to explain how what appears to be a positive message from policy publications becomes bastardised once it reaches the classroom. This, he suggests, is a result of the government being in a rush or looking for a quick fix. Christodoulou (2016) supports this, observing that the government support for the policy was counter-productive and suggesting that the government was too concerned with high-stakes monitoring and tracking rather than low-stakes diagnostics. She advises that the government interpretation of assessment for learning meant it changed from being formative to summative, and in so doing lost the key messages which Black and Wiliam were trying to promote.

Kirby (2013) suggests, however, that Wiliam failed to articulate his ideas clearly, and despite the early criticism of some of his ideas, particularly around some of the aforementioned classroom strategies, has continued to promote these ideas in his later works. Didau (2014) agrees with Kirby and concludes that the 'big idea of AfL is all wrong', and although he concurs that some of Wiliam's ideas such as the traffic light system will not do any harm to the pupils, he also doubts they will have much impact on their metacognition. Wiliam (2011) himself notes that even where teachers have followed his advice, particularly in the case of replacing grades with comments, such comments have not proved to be helpful or formative, with teachers frequently focusing on what is deficient about a piece of work rather than providing helpful comments which would help pupils improve. It could be argued that this was down to a failure to express just what would constitute formative feedback; however, Christodoulou (2016) suggests that the issue may be due to a lack of clear understanding regarding the outcome of a piece of work. She suggests that assessment for learning should be more about how a skill is acquired, with a focus on the method and process of acquiring a skill rather than on the end product. She also advises that while the final outcome may well be summatively assessed, it is important that pupils are supported formatively along the way. For her the processes of formative and summative assessment may well coexist but should also be kept apart where necessary – something which Wiliam failed to acknowledge as he attempted to integrate the two (Wiliam, 2016, cited in Christodoulou, 2016).

Didau (2014) proposes that a major flaw of AfL is the idea that learning can be assessed during the learning process in order to adapt future teaching. He argues that

this is an impossible task for teachers, suggesting that there is no meaningful way of assessing learning during the process. Didau (2014) believes that Wiliam's assertion that learning is being assessed is actually the assessment of performance, and this is problematic since performance is influenced by cues and stimuli from the teacher and does not give an accurate assessment of the retention or transfer of the knowledge and skills taught.

Both Didau (2014) and Kirby (2013) acknowledge that Wiliam's ideas have worth, and as we have seen here it may well be the government's attempts to formalise the process which have proved to be its biggest downfall, since, as Christodoulou (2016: 18) observes, 'the assessment in AfL went from being formative to being summative: no longer assessment for learning but assessment of learning'. Nevertheless, there is still value in applying some of the methods proposed in the *Inside the Black Box* (1998) guidance, which will be discussed in the next section.

APPLYING FORMATIVE ASSESSMENT IN THE CLASSROOM

As we have seen previously, both *Inside the Black Box* (1998) and *Beyond the Black Box* (ARG, 1999) advocated a holistic approach to classroom assessment which was built on formative assessment. The government response to the proposals set out by the ARG was the launch of the Assessment for Learning strategy in 2008, followed by the publication of a set of assessment for learning (AfL) materials and guidance for schools in 2009. These materials were based around the five key elements discussed earlier, offering guidance to schools as to how formative assessment could be implemented in the classroom. The AfL materials offered teachers practical guidance, much of which can be seen echoed in Wiliam's own work, and as discussed previously these can still be seen embedded into practice today.

The assessment for learning strategy outlined five main processes which were key to the successful implementation of a formative assessment programme (see Figure 18.2).

These areas will now be discussed.

QUESTIONING

Questioning serves the purpose of finding out where a learner is at with their learning. However, as previously noted, research from Black and Wiliam (1998a) revealed that the 'wait time' between asking the question and waiting for an answer was around 0.9 seconds, which is not sufficient for a young learner to think about the question and formulate an answer (Black et al., 2003). Rowe (1974) recommends that a more appropriate wait time would be 3 seconds, which she suggests would improve the quality of answers. In the assessment for learning guidance one way of increasing this wait time is the use of mini whiteboards on which pupils will write their answers. Not only does this give pupils time to formulate their answers, it also affords teachers the

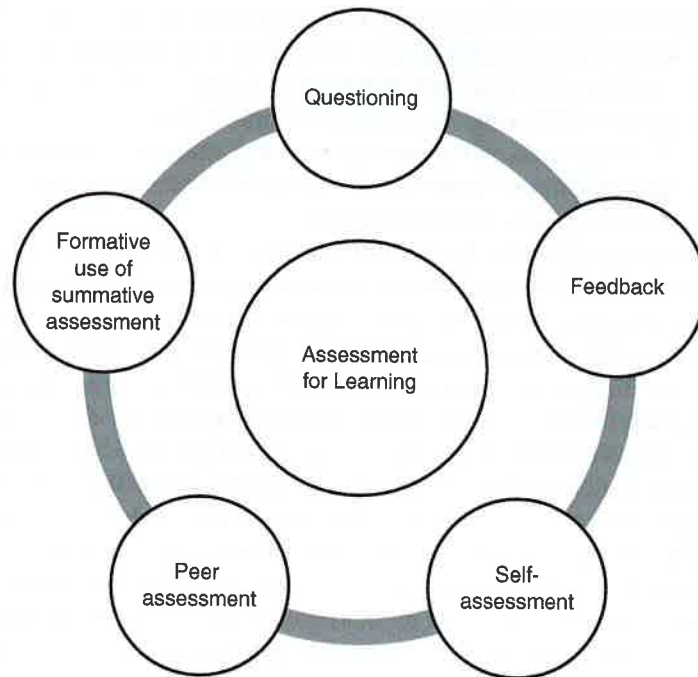


Figure 18.2 Key processes for the successful implementation of AFL

opportunity to assess which students have understood a concept as answers are revealed. A further strategy adopted by some practitioners is the use of 'talk partners' in which pupils share their responses to questions with their partner, thereby giving them the reassurance of the potential accuracy of a response, which in turn can help to build confidence. This is especially effective if talk partners are of a different ability level, allowing discrete opportunities for peer support.

Wiliam (2015) recommends the use of planning for questions mid-session which he refers to as 'hinge questions', arguing that 'lessons never go according to plan, teachers should build plan B into plan A' (2015: 40). He defines the hinge question as a specific type of question designed to ascertain the pupils' understanding of a concept before moving on. Drawing on pupil responses from the hinge question the teacher can decide whether or not to proceed with the lesson or may choose to revisit concepts if the class have not fully grasped the concept. Wiliam (2015) advocates that this should come at a salient point in the lesson, and, in acknowledgement of the challenges of seeking responses from all children in the class, recommends that this can be done through a simple polling exercise.

Wiliam (2015) cautions that designing questions is in itself an art form and observes that 'we have to take great care in designing hinge questions, so students don't get the correct answer for the wrong reason' (2015: 4), arguing that questions which allow

students with the wrong thinking to get the same answers as those with the right thinking are ineffective as a diagnostic tool. Therefore, when planning questions, it is necessary that students with the right thinking and those with the wrong thinking reach different answers. Drawing from Sadler's work, Wiliam proposes that the best way to design such questions is by starting from the students' partial or incomplete understandings and generating questions around these, which then allows the teacher to identify the misconceptions held by students.

A further observation made by Wiliam (2016) is with regard to the traditional method of pupils putting up their hands to answer questions. He advises that some pupils become invisible, with the same pupils always answering the questions, whereas if all pupils are required to respond to questions, inclusivity in the classroom is promoted. Nevertheless, in order to involve all pupils, he warns that participation must be compulsory, with no pupils being given the option of being able to opt out of responding. A method which he employs to encourage participation is the use of lollipop sticks on which he writes the names of all the pupils in the class; the sticks are then selected at random to ensure a non-biased approach to selecting pupils to answer questions. In this way they remain alert because they never know when their name will be selected, and all pupils have an equal opportunity to answer questions. It should be noted, however, that this method comes with a health warning, especially where pupils give the wrong answers. A positive teacher response is paramount here to avoid pupils becoming disheartened and fearful of the process.

A further benefit of questioning is the development of dialogic teaching, in which ongoing talk between teacher and learner is encouraged, largely through the use of open questions. Open questions normally elicit a longer response from a pupil, in which they are able to expand on an answer and reveal much about their thought processes. Further probing by the teacher can reveal the levels of understanding of pupils, but more importantly highlight any misconceptions which can then be addressed. Wiliam (2014b) argues that, by framing the question openly, responses from a range of students can be encouraged at a level appropriate for them. For example, he suggests that two maths problems could be posed asking the pupils which is the most challenging. In this scenario he observes that any ensuing discussion will enable the teacher to ascertain which mathematical issues need to be covered, but more importantly, since the question has been posed in an inclusive way all students can contribute, which supports inclusivity.

FEEDBACK TO STUDENTS

Formative feedback allows both teachers and learners to come together to assess where the learner is at, where they want to be and how they are going to get there. Traditionally feedback comes through teachers marking pupils' work. However, in a 2002 interview Wiliam observed that in marking work teachers spend too long in correcting pupil

errors, especially in cases where the school marking policy sets out that all spelling and grammatical errors must be corrected (Marshall, 2002). Wiliam advises that a learner who sees a page full of corrections will undoubtedly feel demoralised, but more importantly there is nowhere for them to go with this. Instead Wiliam advocates that pupils should be directed towards the error and then given the opportunity to correct this. For example, rather than correcting errors he suggests that the teacher points out how many errors have appeared on the page and gives the pupils time to find these. This in itself has implications for traditional classroom practice in which pupils are rarely given the opportunity to amend mistakes. Wiliam (2011) advocates that pupils should spend twice as much time addressing feedback as a teacher spends giving it, but suggests in reality the opposite is often true. The AfL guidance recommends that marking should be comment-only guidance, avoiding giving any grades or marks. Comments should avoid generalisations such as 'good' or 'well done', and instead teachers should be providing comments which tell pupils what they have done well and the next steps for improvement. More importantly, time should be set aside for reading and addressing comments, and planning should allow opportunities for pupils to take the next steps needed.

SELF-ASSESSMENT

Self-assessment requires that a pupil evaluates their own learning against the learning intentions, and then sets their own goals in terms of what they need to improve and how this might be achieved. However, it should be noted that self-assessment does not come easily to pupils and is something which must be taught. Black et al. observe that 'it is very difficult for students to achieve a learning goal unless they understand a goal and can assess what they need to do to reach it' (2003: 49). The role of the teacher, then, is imperative in supporting this understanding, and strategies for achieving this involve sharing the learning intentions with the learners, and then modelling what a successful outcome might look like. In the classroom this might be seen by displaying learning outcomes on walls or whiteboards, and producing classroom displays with exemplars – sometimes referred to as learning or working walls.

A simple approach to self-assessment proposed by Black et al. (2003) was through the application of a traffic light system, with pupils indicating with red, amber or green labels how far a concept had been understood. In the 2010 BBC2 documentary, *The Classroom Experiment*, Wiliam demonstrated this technique using plastic cups, in which pupils placed the appropriate coloured cup on their desks as they completed their work, thereby giving the teacher instant feedback as to how far they had understood a concept – or, as Wiliam (2016) reflects, allowing teachers to recognise when pupils simply don't get their brilliant lesson. Black et al. (2003) suggest that peer support can also be used here – for example, pupils who indicate a green traffic light might support those with an amber indication, while the teacher concentrates on the red traffic lights.

Guidance from Cambridge International Examinations (CIE) (n.d.) advises that teachers should ask pupils a series of questions about their work to scaffold an approach to self-assessment, and these might include asking them what they understood of the work and what they didn't, how it fits with their existing knowledge and what they might do to improve. This encourages pupils not only to reflect on their work, but also to begin thinking about next steps. Such ownership of their own work can be a powerful motivating force in helping pupils control and direct their own learning.

PEER ASSESSMENT

Peer assessment is a means by which pupils evaluate one another's work, and can be used as complementary to, or even as a prior requirement for, self-assessment (Black et al., 2003). Peer assessment was found to be a motivating force for pupils, with pupils applying more care to their work knowing that their peers would be assessing it. One of the benefits of peer assessment is that pupils are often able to use language that is better understood among themselves, so explanations can be voiced in a more accessible way. Black et al. (2003) also found that pupils were more accepting of critique from peers than from teachers.

As with the renewed attention to marking and feedback it is necessary to provide time in a lesson for peer assessment to proceed, and as with self-assessment it is important to teach pupils to focus on the learning intentions in order that feedback is objective. Nevertheless, if managed appropriately, utilising peer assessment can support pupils in their social skills, and also support their own critical and analytical thinking. At the same time, if pupils are involved in peer assessment then the teacher is able to observe and reflect on what is happening in the classroom to inform future intentions (Black et al., 2003).

Black et al. observe that both peer and self-assessment 'can make unique contributions to the development of students' learning' (2003: 53). However, they also caution that the process does not always come naturally to pupils and reinforces the notion that pupils need to be taught skills in collaboration, listening and turn taking to achieve success. That said, if managed correctly, both peer and self-assessment can strengthen student voice and improve communication between pupils and their teachers (Black et al., 2003).

THE FORMATIVE USE OF SUMMATIVE ASSESSMENT

As we have previously noted, Black and William (1998a) did not disregard the importance of summative assessment, especially since this is the method most widely used to assess where students are with their learning. Indeed, at the time of publishing *Inside the Black Box* it could be argued that high-stakes testing was more prevalent

than ever before as a result of the political commitment to external tests and published league tables. However, Black et al. (2003) advocate that summative testing should work alongside formative assessment, rather than the two being seen as entirely separate processes. Assessment for learning guidance advises returning test papers to pupils so they are able to see where mistakes have been made. For teachers, the analysis of test papers is essential, with a focus on identifying which questions have been answered incorrectly by a majority of pupils and addressing this through future planning.

Black et al. (2003) suggest that for older pupils a reflective review of test papers might help them plan their revision more effectively, so practice or end-of-module assessments can reveal areas of potential weakness and the development of personal revision targets. This can be done through the application of the traffic light system previously discussed, revealing how aspects of formative assessment can be seen working holistically.

It can be stated then that summative assessment should be viewed as a key part of the learning process – not something to be feared, but as a means of charting the process and identifying weaknesses so these can be addressed. Pupils should be the ‘beneficiaries rather than the victims of testing, because testing can help to improve learning’ (Black et al., 2003: 56).

OVERVIEW OF APPLICATION: WILIAM’S IDEAS FOR CLASSROOM PRACTICE

In his text *Embedded Formative Assessment* (2017) Wiliam sets out an intention to ‘provide simple practical ideas about changes that every teacher can make in the classroom to develop his or her practice of teaching’ (2017: 1); as such, Wiliam’s work is designed to be applicable to classroom practice. Drawing from his original Black Box research with Paul Black, Wiliam sets out the five strategies for embedding the principles of formative learning into classroom practice. These strategies will form the basis for this overview in presenting a starting point for the practical application of Wiliam’s ideas. It should be noted that for some of the ideas here time must be taken to develop good practice and pupils will need to be taught some of the strategies before these can be embedded as good practice.

1. Clarifying, sharing and understanding learning intentions and success criteria
 - Share learning intentions verbally as well as displaying these on a board
 - Ensure that learning intentions are framed appropriately for the learner
 - Ensure that learning intentions show cognitive development

(Continued)

- Share the success criteria with learners through clearly explaining what success will look like, for example through providing exemplars
 - Establish that learners have clearly understood what is expected
 - Use hinge questions to monitor achievement against intentions and address any misconceptions in a timely manner
 - As confidence develops encourage learners to identify their own success criteria.
2. Eliciting evidence of learners' achievement
- Find out what the learners already know, for example through the use of well-planned and targeted questions
 - Ensure a sufficient amount of 'wait time' is given to allow learners to think about and prepare a response to a question
 - Establish where ideas come from in order to elicit whether these reflect misconceptions or a different way of seeing things
 - Raise discussion points as an alternative to asking direct questions; this enables the teacher to see how the learners' thinking is developing
 - Use techniques such as traffic light systems and polling to gauge whether learners have understood a concept.
3. Providing feedback that moves learning forward
- Avoid feedback that relies on marks or grades
 - Provide feedback that shows how work can be developed and express clearly why
 - Ensure that feedback reflects learning intentions
 - Avoid feedback that only reflects the negative attributes of a piece of work
 - Prioritise feedback - don't give feedback on all aspects of a piece of work as this can be overwhelming for the learner
 - Ensure the learner can action the feedback, and give them the time to do this
 - Provide feedback that is friendly and personalised.
4. Activating students as instructional resources for one another
- Provide opportunities for cooperative learning
 - Develop a framework for peer tutoring
 - Encourage group cohesion through establishing group goals
 - Develop a class ethos in which students actively seek support from others before asking the teacher - Wiliam (2017) refers to this as C3B4ME (See Three Before Me)

- Allow learners to grade their own and peers' work using specifically designed rubrics
 - Encourage learners to generate their own questions following a unit or topic of work
 - Use group work for test preparation.
5. Activating students as owners of their own learning
- Create an ethos for self-assessment, encouraging learners to assess their own performance against the established success criteria
 - Utilise a traffic light system for learners to identify performance against criteria
 - Open the channels of communication so that learners can express when they haven't understood a concept or where a lesson is moving too fast
 - Use learning journey portfolios so that learners have a visual record of their progress
 - Teach learners how to keep learning logs.

SUMMARY

As we have seen, Black and Wiliam's work on assessment for learning made a significant impact on practices in educational settings, and it could be argued that it was Wiliam's passion for publicly promoting these ideas which led to the success of the Black Box materials. Wiliam has never been afraid of speaking his mind when discussing how improvements to education can be made, and openly expresses that success in schools is not a result of investment in school buildings or decreasing class sizes (Wilby, 2011); instead he promotes the idea of teacher development as being the key to successful learning. His current work focuses on the development of Teacher Learning Communities (TLCs) as a means by which to change teacher behaviour (Goodwin, 2008) through self-supporting self-improvement groups.

It is noteworthy that Wiliam's own relationship with education was not always harmonious, admitting that he was not the best student, and only finding an affinity with education once he had discovered a talent in mathematics. Likewise, his journey into teaching and later academia appeared to be through chance rather than deliberate career moves, starting with a teaching career which began to fund a fledgling music career, and then a research career as a response to not getting a promotion. Nevertheless, despite the route into these careers Wiliam has made a success of all he has done, resulting in a significant number of publications and accolades – testament to his hard work and tenacity.

UNDERSTANDING AND USING EDUCATIONAL THEORIES

William currently works as a freelance academic giving talks and presentations to a range of institutions and dividing his time between the United States, Australia, the United Kingdom and the Far East. He has continued his work on supporting schools in developing their assessment and teaching practices and is also working alongside his wife, Siobhan, on developing their aforementioned vision for building TLCs.

GLOSSARY OF TERMS

Assessment for learning

The process of using assessment as means of informing learning, usually carried out during the learning process.

Assessment of learning

The process of assessing learning at the end of a unit or module, usually conducted through testing.

Black Box

The name derived by Black and William to describe the classroom in which inputs from the outside are fed in with the intention that a set of desirable outputs will follow.

Formative assessment

A means by which student learning is monitored, providing ongoing feedback which can be used constructively by both students and teachers.

Summative assessment

A formal method of testing, usually undertaken at the end of a unit of study, designed to ascertain how far a concept has been understood.

Teacher Learning Communities (TLCs)

A means of personal and professional development in which teachers work in partnership to support one another on self-selected development projects.

FURTHER READING

Bartlett, J. (2015) *Outstanding Assessment for Learning in the Classroom*. Abingdon: Routledge. Practical strategies to support practitioners in making meaningful assessments in the classroom. Specifically focusing on how assessment can be embedded from the outset in the planning process.

Chiles, M. (2020) *The Craft of Assessment: A whole school approach to assessment for learning*. Woodbridge: John Catt Educational.

Suitable for both primary and secondary teachers, this book provides practical solutions to developing a whole-school approach to creating a climate conducive to effective assessment. The author focuses on the five principles of condensing, reflecting, assessing, feed forward and target-driven improvement.

Clarke, S. (2001) *Unlocking Formative Assessment*. Didcot: Hodder Education.

Practical suggestions as to how formative assessment can be used in the classroom, unpacking some of the key terms proposed by Black and Wiliam.

Clarke, S. (2005) *Formative Assessment in Action: Weaving the elements together*. Didcot: Hodder Education.

Informed through action research, this book brings together key aspects of formative assessments linked to lesson plans and exemplars, showing how these can be successfully applied to practice.

Jones, K. (2021) *Wiliam's and Leahy's Five Formative Assessment Strategies in Action*. Woodbridge: John Catt Educational.

Building on the work of Wiliam and Leahy, this book examines five formative assessment strategies and how they can be applied in the classroom, utilising case studies to show how formative assessment can be successfully embedded into the curriculum.

Wiliam, D. (2016) *Leadership for Teacher Learning: Creating a culture where all teachers improve so that all pupils succeed*. West Palm Beach, FL: Learning Sciences International.

Research-informed guidance as to how student achievement can be increased through the creation of a structured and rigorous learning environment.

REFERENCES

- Assessment Reform Group (1999) *Assessment for Learning: Beyond the Black Box*. Cambridge: University of Cambridge School of Education.
- BBC (2010) *The Classroom Experiment*. Available from: www.bbc.co.uk/programmes/b00txzwp [accessed 12 March 2018].
- Black, P. and Wiliam, D. (1998a) *Inside the Black Box*. Brentford: Department of Education and Professional Studies, GL Assessment.
- Black, P. and Wiliam, D. (1998b) Assessment and classroom learning. *Assessment in education: Principles, Policy and Practice*, 5(1), 7–74.
- Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. (2003) *Assessment for Learning: Putting it into practice*. Maidenhead: Oxford University Press.
- Bloom, B.S. (1969) Some theoretical issues relating to educational evaluation. In: Tyler, R.W. (ed.) *Educational Evaluation: New roles, new means: The 68th yearbook of the National Society for the Study of Education (Part II)* (Vol. 68(2)). Chicago, IL: University of Chicago Press.
- Butler, R. (1988) Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, 58, 1–14.
- Cambridge International Examinations Teaching and Learning Team (n.d.) *Getting started with assessment for learning*. Available from: www.cambridge-community.org.uk/professional-development/gswaf/index.html [accessed 12 March 2018].

- Christodoulou, D. (2016) *Making Good Progress? The future of assessment for learning*. Oxford: Oxford University Press.
- Department for Children, Schools and Families (2009) *Getting to Grips with Pupil Progress*. London: HMSO.
- Department for Education (2018) *Assessment without Levels: Qualitative research. Research report*. London: HMSO.
- Didau, D. (2014) *Why AfL might be wrong and what to do about it*. Available from: www.learningspy.co.uk/myths/afl-might-wrong/ [accessed 12 March 2018].
- Goodwin, G. (2008) *Dylan Wiliam: Profile interview with Gail Goodwin*. Slough: NFER. Available from: www.dylanwiliam.org/Dylan_Wiliams_website/Bios_files/Gail%20Goodwin%20%28NFER%29.doc [accessed 12 March 2018].
- Kirby, J. (2013) What can we learn from Dylan Wiliam and AfL? *Pragmatic Education*. Available from: <https://pragmaticreform.wordpress.com/2013/03/30/afl/> [accessed 12 March 2018].
- Marshall, B. (2002) Thinking through assessment: An interview with Dylan Wiliam. *English in Education*, 36(3), 47–60.
- McInerney, L. (2015) *Profiles: Dylan Wiliam, Emeritus Professor of Educational Assessment*. Institute of Education. Available from: <https://schoolweek.co.uk/dylan-wiliam/> [accessed 12 March 2018].
- Menzies, L. (2010) *Policy diffraction*. Available from: www.lkmco.org/policy-diffraction/ [accessed 12 March 2018].
- Rowe, M.B. (1974) Wait-time and rewards as instructional variables, their influence on language, logic, and fate control. *Journal of Research in Science Teaching*, 11(2), 81–94.
- Sadler, D.R. (1989) Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–44.
- Scriven, M. (1967) The methodology of evaluation. In: Tyler, R.W., Gagné, R.M. and Scriven, M. (eds) *Perspectives of Curriculum Evaluation 1*. Chicago, IL: Rand McNally.
- Wilby, P. (2011) Teaching guru is optimistic about education. *The Guardian*, 18 January. Available from: www.theguardian.com/education/2011/jan/18/teaching-methods-government-reforms [accessed 12 March 2018].
- Wiliam, D. (2011) *Embedded Formative Assessment*. Bloomington, IN: Solution Tree Press.
- Wiliam, D. (2014a) *Formative assessment and contingency in the regulation of learning processes*. Paper presented in symposium 'Toward a Theory of Classroom Assessment as the Regulation of Learning', Annual Meeting of the American Educational Research Association, Philadelphia, PA, April.
- Wiliam, D. (2014b) The right questions, the right way. *Educational Leadership*, 7(6), 16–19. (Association for Supervision and Curriculum Development).
- Wiliam, D. (2015) Designing great hinge questions. *Educational Leadership*, 73(1), 40–44. (Association for Supervision and Curriculum Development).
- Wiliam, D. (2016) The secret of effective feedback. *Educational Leadership*, 73(7), 10–15. (Association for Supervision and Curriculum Development).
- Wiliam, D. (2017) *Embedded Formative Assessment*. Bloomington, IN: Solution Tree Press.