



Assessment for learning: effects and impact

María Teresa Flórez and Pamela Sammons
Oxford University Department of Education

School improvement: international reviews of best practice

Working with partners including the Department of Education at Oxford University, the Centre for Equity in Education at the University of Manchester, the University of Glasgow, the University of Nottingham and the Hong Kong Institute of Education, CfBT Education Trust has commissioned a series of reviews of international literature. These reviews cover a range of topics related to school improvement including assessment for learning; the inclusion of students with special educational needs; effective teaching practice; school self-evaluation; and successful school leadership.

The idea that schools can impact positively on student outcomes is a crucial driver in the rise of interest in school improvement research and practice. These reviews highlight international examples of best practice in order to effect change and identify how effective school improvement manifests itself. It forms a useful tool for schools and school leaders, but also acts as a lesson for policymakers in terms of what works around the world.

This review focuses on: **Assessment for learning**

Assessment for learning – where the first priority is to promote learning – is a key means of initiating improvement. The features, strategies and principles underpinning assessment for learning form the basis of this review.

The other four reviews in this series focus on:

From exclusion to inclusion

With a specific focus on children with special educational needs (SEN), this review addresses the forms of classroom practice that can help all children to participate. The review particularly focuses on elements of inclusive education and the implications for schools and school leaders.

Effective teaching

Teachers are one of the key elements in any school and effective teaching is one of the key propellers for school improvement. This review is concerned with how to define a teacher's effectiveness and what makes an effective teacher. It draws out implications for policymakers in education and for improving classroom practice.

School self-evaluation for school improvement

School self-evaluation can be a fundamental force in achieving school improvement. This review establishes what the key debates are in relation to school self-evaluation, what principles and processes are associated with it, and what the implications are for school self-evaluation as a means of leading school improvement. The review incorporates a framework for conducting self-evaluation and case study examples from systems and schools that have previously undergone the process.

Successful leadership

School leaders are under considerable pressure to demonstrate the contribution of their work to school improvement, which has resulted in the creation of a wide range of literature which addresses leadership in the context of school improvement. This review pays particular attention to issues including transformational leadership, instructional/pedagogical leadership and distributed leadership.

CfBT is a world authority on school improvement. We work directly with schools and governments improving education outcomes through evaluation, training and professional development programmes. This series of reviews fits into our aim to develop evidence for education and supports our goal to provide school improvement programmes which are evidence based.



Contents

Executive summary	2
Introduction	3
What is assessment for learning?	6
Assessment for learning put into practice	12
Effects and impact of assessment for learning	16
Guidelines for policymaking and practice	19
Glossary	23
References and further reading	24

Executive summary

Assessment that is *for* learning, as opposed to merely *of* learning, looks forward as well as back. Teachers who assess in this way are concerned not just to confirm and verify what their students have learnt, but also to help their students and themselves understand what the next steps in learning should be and how they might be attempted. This kind of assessment has a ‘formative’ purpose: it helps to shape what lies ahead rather than simply to gauge and record past achievements.

The main strategies considered important for Assessment for Learning (AfL) – sharing learning goals, formative feedback, peer and self-assessment, and the formative use of summative tests – have been found to be overwhelmingly positive in terms of their potential to promote improvements in teachers’ classroom practice.

This review proposes that in order to encourage AfL, subject departments:

- have an atmosphere in which teachers are expected to watch others in action – to actively support peer observation
- recognise and value current skills and help teachers to identify their current formative practice.
- have meetings where teachers discuss learning
- give teachers time to plan well by encouraging them to mark less, but mark better.

The review concludes by identifying two sets of guidelines, for policymaking and practice. The former include:

- the use of national testing and assessment strategies and the role of public examinations must be evaluated in terms of their consistency with AfL, implementing if necessary national assessment systems with a greater degree of teacher involvement
- policymakers should give clear information to the educational community and the public about the views on assessment which are to be promoted or prioritised in an education system, in order to avoid confusion and overlapping of approaches among practitioners.

The latter include:

- whole-school commitment to the AfL approach – involvement and leadership of senior staff is especially important to embed AfL
- clarity of messages is also relevant at the school level – a change towards assessment for learning should be informed and explained to all stakeholders, including students and parents
- teachers should attend well developed and recognised professional development programmes that help them to develop a flexible and deep understanding of the sense of the approach, and not just to accumulate a set of techniques they might apply without the required level of reflection.

Introduction

The two main aims of this paper are to give an outline of the main characteristics of Assessment for learning (AfL) and to synthesise key research evidence on the effects and impact of this approach through a systematic process of literature review.

The paper provides a brief description of the main features, principles and strategies that underpin the concept of assessment for learning, as well as information on how this approach has been put into practice in different contexts. Findings about the effects and impact of assessment for learning on student outcomes and teachers' classroom practice are presented and some guidelines for policy-making and practice on assessment derived from research are developed. In addition, the paper identifies some of the criticisms that have been made of this perspective and the limitations of current knowledge on its impact.

The most quoted definition of AfL is the one developed by the Assessment Reform Group (ARG) in the UK:

'Assessment for learning is any assessment for which the first priority in its design and practice is to serve the purpose of promoting pupils' learning. It thus differs from assessment designed primarily to serve the purposes of accountability, or of ranking, or of certifying competence.'

An assessment activity can help learning if it provides information to be used as feedback by teachers, and by their pupils in assessing themselves and each other, to modify the teaching and learning activities in which they are engaged. Such assessment becomes 'formative assessment' when the evidence is actually used to adapt the teaching work to meet learning needs.'¹

AfL originated as a response to the need for assessment more consistent with current pedagogical tendencies that link with constructivist – rather than more traditional – approaches to teaching and learning. AfL goes beyond more psychometric and behaviouristic traditions of assessment that focus on measuring individual students' performance in specific domains against externally norm-referenced distributions of attainment.

¹ Black et al. (2004: 2-3). The Assessment Reform Group originated in 1989 as a voluntary group of researchers concerned with providing a research basis for decisions on assessment policymaking and practice in the UK. Their work has been closely related to teachers and educational practice in order to complement assessment theory with the needs and the wisdom of practice. Most of the texts selected for this review consider the ARG's conceptualisation of assessment for learning, either as their unique definition of the concept or intertwined and contrasted with the ones proposed by other authors.

Based on the notion of formative assessment, the ARG has sought to develop a 'translation' of this perspective into practice through the development and study of joint work with teachers that seeks to improve the learning processes for students and teachers.

As a result of this experience, which has been adopted and adapted in multiple contexts, they have constructed some fundamental principles that comprise a framework for assessment for learning. These principles state that assessment for learning:²

- is part of effective planning
- is central to classroom practice
- promotes understanding of goals and criteria
- is sensitive and constructive
- fosters motivation
- recognises all educational achievement
- focuses on how pupils learn
- helps learners know how to improve
- develops the capacity for peer and self-assessment
- is a key professional skill.

With these principles as the framework, some strategies have been developed and are claimed to have a number of positive effects on students and teachers. The most commonly-mentioned features of AfL in the literature are the better use of questioning, feedback, peer and self-assessment and the formative use of summative tests where these are used in the schools or system concerned. Some research and inspection evidence has also been provided regarding the impact of AfL on students' achievement, though this aspect requires further rigorous research to arrive at stronger conclusions.

Although most of the literature evaluates the effects of AfL as positive, contextual aspects emerge as possible obstacles for the feasibility of the approach, especially in those contexts in which the ideal conditions observed in research are not given. Further research on this is required as well, but some guidelines can be given on the basis of currently available studies.

Policymakers should consider a careful design of dissemination strategies and possible contradictions between different policies. They should give clear messages to the educational community and the public about the view of assessment which is to be promoted or prioritised; the provision of support for dissemination processes through fostering school leadership; enough flexibility to allow some level of appropriation by practitioners; processes of monitoring the progress of dissemination; and commitment to sustain the policy over time.

A whole-school commitment is required if AfL is to be successfully adopted, with all stakeholders being informed about and involved in the process. In addition, senior staff can play an especially important role in promoting, monitoring and generating the right conditions for adequate dissemination and implementation of AfL in their school through encouraging and facilitating teacher collaboration and discussion of the way AfL can inform and develop their current classroom practice.

² ARG (2002).

Review methodology

The following review question guided the search process:

What does recent research indicate about the impact of teachers' use of assessment for learning, rather than merely of learning?

Relevant databases in English and Spanish were searched (ERIC, AUEI, BREI, COPAC, CERUK, SCIELO, UNESCO, among others), employing keywords centred on the use of assessment for improving learning. From the initial 478 sources found, three phases of data collection were carried out according to a series of inclusion/exclusion criteria related to type of document, context of research, relevance of research and date. Repetition of research projects and their findings was also avoided and the most recent publications were prioritised.

As a result of this data collection process 33 texts were considered of high relevance for the review. Six texts were selected for their overarching character, that is, because they were systematic reviews, meta-analyses, summaries of research or methodical analyses, about results of research on assessment for learning (AfL) in different contexts. Another 27 articles and chapters presenting research evidence of individual projects were selected based on their relevance in relation to the purposes of this review. In brief, this selection represents evidence about AfL developed from 1998 to 2010 in a great variety of school contexts around the world.

What is assessment for learning?

Principles

The ARG defined ten principles in order to summarise the main aspects of the frame in which any practice of AfL should be understood and developed:

<ol style="list-style-type: none"> 1. AfL is part of effective planning 2. AfL is central to classroom practice 3. AfL promotes understanding of goals and criteria 	<p>These first three principles can be understood as a whole. In brief, they refer to the need to recognise assessment not as a mere accessory to pedagogical practice, but as an integral part of it. Assessment must be intertwined with all the moments of a learning process and, thus, must be considered when planning. For this to happen, teachers must define clear learning goals or criteria and be able to share them with students in an understandable way. Along with this, students should be constantly reminded of these criteria or learning goals during the learning process, their learning evaluated and feedback given to analyse the progress of students and take decisions according to this evidence.</p>
<ol style="list-style-type: none"> 4. AfL is sensitive and constructive 5. AfL fosters motivation 6. AfL recognises all educational achievement 	<p>This second set of principles is related to the impact of assessment in shaping students' motivation, especially in terms of the nature of the feedback they receive. Teachers should be careful in what they say to students and try to give descriptive feedback exclusively centred on the quality and content of each student's work rather than use value-laden terms such as 'good' or 'poor'. They should also suggest ways for students to improve their work. In the context of AfL, there is not only an excellence level which all must achieve in order to have recognition; any learning progress made by the student in relation to his or her previous state deserves recognition and positive feedback.</p>
<ol style="list-style-type: none"> 7. AfL focuses on how pupils learn 8. AfL helps learners know how to improve 9. AfL develops the capacity for peer and self-assessment 	<p>The process through which students learn must be a focus of attention in classroom practice, both for teachers and students. This involves developing awareness in the student about his or her learning processes, and increasing autonomy through practices of peer and self-assessment in order to support students in developing their own responsibility for their learning. Giving feedback to students on how to improve, and not just on their mistakes, also contributes to the development of autonomous thinking and learning.</p>
<ol style="list-style-type: none"> 10. AfL is a key professional skill 	<p>This principle highlights the complexity involved in taking assessment for learning into practice, as it requires teachers to learn how to work from this perspective and to develop the necessary skills for doing so. The ARG recognises here the need for good quality professional development programmes as a fundamental requirement for the successful implementation of assessment for learning in classroom practice.</p>

Strategies

With these overarching principles as a framework, research on AfL considers different strategies to take the approach into practice.

Key idea: All the studies and reports reviewed concentrate on at least one of the same four aspects that are understood as characteristic of AfL:

- Questioning
- Feedback
- Peer and self-assessment
- The formative use of summative assessment.

The original source for these aspects is the ARG,³ which confirms that most recent research on assessment for learning is based on the work of these authors. The four strategies and the key research findings for each one of them are described below.

The role and importance of questioning

The importance of teachers using appropriate questioning to support student learning is an idea that emerges from the initial study of the ARG. Here the researchers noticed some problems in the way in which teachers were formulating and using questions during class interaction. They concluded that the questions used by teachers were often insufficiently challenging for the students. In addition, the time given to elaborate on an answer was too short.⁴ It was discovered that a better practice in questioning was essential for putting the principles of AfL into practice. The specific modifications were related to an increase in the 'wait time' (time given for students to think of an answer)⁵ and an improvement in the quality of questions employed by teachers in terms of greater use of those that are open rather than closed in character. This encourages students to use higher-order thinking skills to answer questions.⁶

Such practice requires the generation of a non-threatening climate of dialogue, where students feel free to share their ideas and expose their misconceptions without fear of giving an incorrect answer. The idea is that student learning is enhanced by opportunities for discussion of misconceptions and clarification of these in the classroom so that all students can learn from one another's mistakes or misconceptions.⁷

³ Black et al. (2010).

⁴ These conclusions are in line with those of many studies of effective teaching that draw attention to the need for teachers to provide challenging interactive teaching and sufficient 'wait' time to encourage students to answer (see Ko et al., 2013).

⁵ Changes in time for questioning were also considered by Condie et al. (2005); Kirton et al. (2007); Gipps et al. (2005); Webb and Jones (2009); Kellard et al. (2008).

⁶ Improvements in the quality of questions are also recognised as necessary by Torrance and Pryor (2001); MacPhail and Halbert (2010); DfES (2007); Condie et al. (2005); Dori (2003); Kirton et al. (2007); Tapan (2001); Gipps et al. (2005); Webb and Jones (2009); Gioka (2006); Kellard et al. (2008); Stiggins and Arter (2002).

⁷ Torrance and Pryor (2001); Willis (2008); Kirton et al. (2007); and Tapan (2001) refer to the need of such an environment in order to promote good questioning.

The importance of feedback

Frequently mentioned in the literature, good quality feedback is viewed as a fundamental aspect in the use of AfL.⁸ This idea is based on a study that concluded that students who were given *comment-only marking** and no score on their work later obtained results in achievement tests that were on average one standard deviation higher than the students who had received only marking-centred feedback.⁹

Informative and descriptive feedback* is held to be more appropriate for a learning-centred perspective than the exclusive practice of marking work as right or wrong and giving just an overall mark. The latter is deemed to promote comparison and competition between students, and may in consequence damage their self-esteem if they get low scores compared with other students in their class. Research on three empirical studies in Germany demonstrated that individual reference has more positive effects on students than social reference as assessment norms for teachers.¹⁰ Through individual reference and feedback students are encouraged to concentrate on reviewing their own progress and learning process instead of situating themselves in terms of notions of ability as either low or high achievers in comparison with their peers.

Qualitative studies provide evidence of the positive effect of feedback according to teachers', students' and other practitioners' perspectives. Such studies indicate that formative feedback is generally recognised as a tool for:

- improving learning¹¹
- changing traditional classroom practices¹²
- clarifying learning criteria¹³
- raising achievement.

For example, research has shown that school staff identify feedback as one of the more influential aspects in raising students' scores in Key Stage 3 national assessment results in England.¹⁴

It should be mentioned that the positive role of formative feedback is also recognised in other approaches to effective teaching including interactive whole class teaching and direct instruction.¹⁵

Some authors have criticised the research evidence on AfL due to a limitation in the methodical analysis: the studies do not tend to consider the influence of this factor (feedback) in isolation from others.¹⁶ Thus, it is hard to establish the specific effect of feedback separated from other potential influences on students' learning outcomes, such as questioning. Thus, reviews of effective teaching suggest that feedback is important but also highlight the role of other important aspects of teachers' classroom behaviour and practices. It is suggested that formative feedback needs to be seen as part of a repertoire of effective classroom practices.¹⁷

⁸ Along with the six overarching sources considered for this review (systematic reviews and methodological analyses), feedback as a relevant aspect of AfL is recognised by Torrance and Pryor (2001); MacPhail and Halbert (2010); DfES (2007); Weeden and Winter (1999); Brookhart (2001); Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Kirton et al. (2007); Tapan (2001); Gipps et al. (2005); Webb and Jones (2009); Bartoochi and Keshavarz (2002); McDonald and Boud (2007); Carless (2005); Gioka (2006); Kellard et al. (2008); Stiggins and Arter (2002); Klenowski (2009). For a recent study of feedback which includes student teachers' and students' perceptions of it, see also Hargreaves (2011).

* See Glossary.

⁹ Butler (1987).

* See Glossary.

¹⁰ Köller (2005).

¹¹ Carless (2005).

¹² Kellard et al. (2008); Gipps (2005).

¹³ Torrance and Pryor (2001).

¹⁴ DfES (2007: 31).

¹⁵ Ko et al. (2013).

¹⁶ Tierney and Charland (2007).

¹⁷ Ko et al. (2013).

To understand the notion of feedback in the context of assessment for learning, a crucial factor to consider is the need for feedback to be closely linked with the use of explicit and shared learning goals or criteria, as many authors recognise in the sources reviewed.¹⁸

It is argued that students can only understand the feedback they receive from teachers or peers if they are first clear about the purposes of learning and what is expected from them. This is the necessary platform that should underpin constructive feedback. Advocates claim that the feedback comments should always go back to these initial criteria in order for both students and teachers to share a developing understanding about the progress that has been made towards them and to identify strategies that may help to close the learning 'gap' between actual achievement and the desired learning goals.

On the basis of these criteria, feedback should give students some guidelines not only about what is incorrect in their work, but also on what has been done correctly according to the initial learning goals or expectations. In addition, and perhaps most important, explicit guidance on what they could do to improve and keep on progressing towards expectations should also be given. This feedback practice focused on improving work is seen as crucial to the learning process and is believed to promote greater student motivation and commitment to enhancing their own learning than assessments that mark work only in a summative way as either right or wrong and then apply an overall mark.

A literature review of 39 studies from Australia, New Zealand, Asia-Pacific and the US, concluded that there are two kinds of feedback: one that moves from students' work to teachers and one that circulates from teachers to students.¹⁹ It is argued that learning results from their alternation.²⁰ Regarding the first kind of feedback, UK researchers²¹ indicate that the exercise of writing comments in students' work also has a major benefit in *enhancing teachers' own professional learning*. This is because it helps teachers to evaluate the quality of the learning tasks they have designed. By studying these in relation to examples of students' work during the process of providing written feedback they can establish whether some learning activities or tasks are better able to scaffold students' learning. This process enables them to distinguish which tasks or activities prove to be good instruments for obtaining appropriate evidence of student learning in relation to some specific learning goals or criteria.

Various authors in the AfL field draw attention to the existence of multiple types, tools and methods for giving feedback to students.²² Types include evaluative feedback in contrast to descriptive or informative feedback and a distinction between oral and written feedback. Among the methods and tools, authors mention: two stars and a wish, now and next steps, traffic-lighting, comment-only marking, comments linked to criteria, review at the end of the lesson, formative use of summative tests, progress reports, reflective comments in a portfolio, a teacher register of students' presentations, feedback through peer assessment, computer-generated feedback and the use of rubrics.*

¹⁸ Black et al. (2010); Tierney and Charland (2007); Torrance and Pryor (2001); MacPhail and Halbert (2010); DfES (2007); Weeden and Winter (1999); Condie et al. (2005); Hayward and Spencer (2010); Kirton et al. (2007); Gipps et al. (2005); Webb and Jones (2009); Bartoochi and Keshavarz (2002) / as learning goals; McDonald and Boud (2007); Carless (2005); Gioka (2006); Kellard et al. (2008); Stiggins and Arter (2002); Klenowski (2009). However, it must be remembered that providing clear and explicit learning goals and reviewing these are also features of effective teaching practices including direct instruction (Rowe, 2006; Ko et al., 2013).

¹⁹ Hodgson and Pyle (2010).

²⁰ This two-way notion of feedback is also highlighted in the study of Webb and Jones (2009) and in French literature on assessment according to Allal and Mottier's review (2005).

²¹ Black et al. (2010).

²² DfES (2007); Weeden and Winter (1999); Condie et al. (2005); Hayward and Spencer (2010); Kirton et al. (2007); Tapan (2001); Gipps et al. (2005); Webb and Jones (2009); Bartoochi and Keshavarz (2002); Carless (2005); Gioka (2006); Klenowski (2009); Tierney and Charland (2007).

* See Glossary.

Nonetheless, research by AfL advocates²³ refers to some common concerns articulated by teachers about the use of comment-only feedback. Not only is this viewed adversely by some because it is very time consuming, but also because of its perceived inconsistency with the administrative requirements of schools and the feedback parents are used to receive. In connection with the problem of time, the authors mention again the potential of enhanced collaboration between colleagues as a crucial aspect that can help to diminish the workload by sharing experiences and ideas on how to make this work more manageable. In relation to possible differences between the expectations and requirements of school administrators and parents compared with the new AfL practices on feedback, the group of researchers refer to teachers' surprise when none of these stakeholders complained about this new way of giving feedback.²⁴ However, the ideal conditions of this research must be taken into account in this aspect, as staff in these schools were volunteers willing to learn about AfL, which is not always the case, as the Ofsted report based on inspection evidence in England demonstrates.²⁵

Peer and self-assessment

These are recommended by some authors as good strategies for students to develop greater autonomy in relation to their own learning²⁶ along with promoting their metacognition²⁷ and collaborative learning skills.²⁸ Some researchers²⁹ again highlight the difficulty of this, as they find that most students are not accustomed to thinking about their work in relation to particular learning goals and argue that it takes time and training to develop in students the required level of awareness to be able to understand their own learning process and act on it. This requires, again, the definition of criteria that need to be communicated to students in a transparent way. In the absence of these criteria, students will not have a basis on which to support judgements about their work or that of their peers.

Peer assessment is regarded as particularly valuable in terms of providing additional formative feedback. It is argued that students demonstrate a greater acceptance of observations about their work when they come from a peer rather than from a teacher. The use of a similar language by their peers also contributes to making feedback more accessible and understandable. This requires, however, the previous development, among the student body, of learning how to work well together in a group.³⁰

Again, different methods are developed in the context of research, with the purpose of facilitating these practices through a concrete procedure.³¹ The most commonly mentioned are self-marking, paired marking, plenary self-evaluation, self-assessment journals, traffic-lighting, question setting, comment-only marking, concept mapping, colouring squares for goal statements and jigsawing.* However, criticisms have been made³² that highlight the lack of research evidence to indicate the relative merit of each of these methods in promoting better student outcomes.

²³ Black et al. (2010).

²⁴ Black et al. (2010).

²⁵ Ofsted (2008). This concern with research being based on ideal conditions which are not the ones observed in current practice is mentioned in the studies and reports of Torrance and Pryor (2001); DfES (2007); Dori (2003) and McDonald and Boud (2007).

²⁶ Students' involvement in general and self- and peer-assessment in particular are recognised as valuable aspects to improve classroom practice and students' learning in the studies and reports of Torrance and Pryor (2001); MacPhail and Halbert (2010); DfES (2007); Weeden and Winter (1999); Brookhart (2001); Condie et al. (2005); Dori (2003); Kirton et al. (2007); Tapan (2001); Gipps et al. (2005); Webb and Jones (2009); Bartoochi and Keshavarz (2002); McDonald and Boud (2007); Carless (2005); Gioka (2006); Kellard et al. (2008); Stiggins and Arter (2002); Klenowski (2009).

²⁷ Brookhart (2001); Gipps et al. (2005); Stiggins and Arter (2002).

²⁸ Kirton et al. (2007); Webb and Jones (2009); Stiggins and Arter (2002).

²⁹ Black et al. (2010).

³⁰ Kirton et al. (2007); Black et al. (2010).

³¹ Hodgson and Pyle (2010); Thompson and Wiliam (2008).

* See Glossary.

³² Tierney and Charland (2007).

Some other gaps in the AfL research evidence are related to possible risks and contextual obstacles that might impede students' involvement, and to the relationship between 'fostering autonomy and encouraging constructive collaboration'.³³ Among the risks, it is frequently mentioned that some teachers show resistance to encouraging greater student involvement and participation as this might involve changing the traditional power relationships in the classroom, with teachers losing part of the control of their classroom situation they hold in more traditional teaching arrangements as a consequence.³⁴

The formative use of summative assessment

There are different perspectives about the relationship between formative and summative assessment. Some authors think of them as two separate and hardly reconcilable faces of assessment³⁵ while others feel they can be complementary and that summative assessment results can also be helpful as part of the formative process.³⁶

This last perspective emerged from teachers' practice during the research initiative of the ARG to implement AfL. This finding challenged the initial idea of the ARG that formative and summative assessment should be viewed as two different processes.³⁷ This division was perceived to be artificial by most teachers involved in the research, because they felt that both processes were integrated and conducted as a normal part of their daily practice. Due to this finding, ways of harmonising the two approaches began to be explored in a joint effort of teachers and researchers involved in the project and have been explored in further studies.

Such harmonisation and integration has been developed through practices such as fostering students' involvement in reviews of summative assessments; preparing for exams by asking students to formulate their own questions, asking them to discuss these with one another; reviewing examination results collectively; peer assessment of summative assessments; and traffic-lighting.* However, this reconciliation was only possible with summative assessments over which teachers had some control. The ARG believes that high-stakes national testing can operate as a barrier to the feasibility of introducing AfL as a main approach to developing teachers' classroom practice because of its links with accountability mechanisms. This is because a strong emphasis on high-stakes tests may distort teaching, which becomes too much focused on the requirements of the test at the expense of promoting deep learning by students.³⁸ Concerns of this kind, related to the need of teachers to reconcile formative and summative assessment, have emerged in different studies as one of the main difficulties for the implementation of AfL.³⁹

It is also important to note that a US study has indicated that this integration of summative and formative assessment is an important strategy for successful students. The results of this research show that such students commonly use summative assessment information for improving their learning as well as for judging their achievement at a specific point in time.⁴⁰

³³ Tierney and Charland (2007: 16).

³⁴ MacPhail and Halbert (2010); DFES (2007); Willis (2008); Kirton et al. (2007); Stiggins and Arter (2002). See also Adamson (2011).

³⁵ That is the case in the studies of Kirton et al. (2007); McDonald and Boud (2007); Carless (2005) and Gioka (2006).

³⁶ Torrance and Pryor (2001); Brookhart (2001); Condie et al. (2005); Tapan (2001); Webb and Jones (2009); Stiggins and Arter (2002). See also Tam and Lu (2011).

³⁷ Black et al. (2010).

* See Glossary.

³⁸ Black et al. (2010); Black et al. (2004).

³⁹ Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Kirton et al. (2007); Gipps et al. (2005); Kellaghan (2004).

⁴⁰ Brookhart (2001).

Assessment for learning put into practice

There seems to be a considerable global interest and recognition of the need for a formative perspective on assessment, despite the difficulties such a perspective encounters in each context. AfL – or at least some of the practices related to it – has been promoted, piloted and/or implemented in a wide variety of contexts, such as the UK, Australia, New Zealand, the US, Barbados, Canada, Israel, Portugal, Belgium, Hong Kong, Iran, Chile, Singapore, some Francophone African countries and the Netherlands.⁴¹

Most research has been related to specific professional development programmes, based on developing a close collaboration between teachers and researchers.⁴²

These programmes are usually structured to include:

- some introductory and some theoretical sessions with the whole group of teachers and academics/researchers
- a series of more personalised sessions with small groups of teachers, sometimes with the presence of one member of the research staff
- in some cases, a third strand of activity involving class observation which includes moments of discussion between the researcher and teacher before and/or after the session.

The aspect in which such professional development programmes vary most is the degree to which they guide teachers' work. Just to give an idea, it is possible to find approaches in which teachers and researchers collaboratively construct and suggest ways of implementing AfL, as in the study of the ARG,⁴³ and others in which teachers are given some basic principles and a list of techniques to implement the approach, with researchers expecting appropriation of it in practice.⁴⁴

The following boxes illustrate two cases of programmes with different degrees of guidance and participation of teachers. The kinds of studies carried out in the context of these programmes are also quite varied in their approach, with a predominance of qualitative methodologies.⁴⁵

Example 1: teachers highly guided by researchers

'Formal training in self-assessment skills was undertaken using twelve modules designed by the researcher (...). The rationale for the training modules was based on the professional experience of the researcher while working in high schools in the West Indies for over 30 years. (...) Conceptually, the modules focused on constructing, validating, applying and evaluating criteria to apply to students' work. The modules encouraged students to practise naming assessment features and giving reasons for assessing the worth of their work. (...) Teachers were first trained as a group to use the modules through a series of interactive workshops with the researcher offering simulations of anticipated classroom situations. Feedback and continuous interaction among the teachers continued throughout the duration of the programme, promoting group cohesion and allowing for information exchange.⁴⁶

⁴¹ Tierney and Charland (2007); Hodgson and Pyle (2010); Azúa and Bick (2009); Black et al. (2010); Kellaghan (2004); Tan (2011).

⁴² Tierney and Charland (2007); Torrance and Pryor (2001); MacPhail and Halbert (2010); DFES (2007); Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Kirton et al. (2007); Webb and Jones (2009); Carless (2005); Kellard et al. (2008); Azúa and Bick (2009); Tapan (2001); Bartoochi and Keshavarz (2002); McDonald and Boud (2007).

⁴³ Black et al. (2004) and (2010); William et al. (2004).

⁴⁴ This is the case in the study of Thompson and William (2008).

⁴⁵ Tierney and Charland (2007).

⁴⁶ McDonald and Boud (2007: 213-4).

Example 2: teachers more actively engaged

'The Education Department of Jersey invited a team from King's College London to lead a professional development programme on formative assessment for Jersey teachers. The professional development programme followed similar principles and built on methods used previously (Black and Wiliam 2003). The (...) initiative aimed to enable the lead teachers, who were chosen by their head teachers, to develop their own classroom practice by experimenting with formative assessment techniques and ideas and by evaluating the change through reflection and discussion with the King's team and with each other. This was facilitated by workshops, lesson observations and discussions. (...) As the lead teachers developed their own practice, they were expected to encourage experimentation and to share their experiences with other teachers in their own and other schools, especially those involved in subsequent phases of the project.⁴⁷

Such predominance has allowed researchers to place a strong emphasis on involving the teachers' voice and exploring their opinions and reactions to the development programme in these studies. Nonetheless, some authors highlight the general absence of students' perspective about this new way of understanding assessment and whether students find it helpful.⁴⁸ Thus, although many claims are made that AfL approaches boost student learning, motivation and engagement, only a few studies have considered the pupils' perspective explicitly.⁴⁹ In general, all these studies demonstrate in different degrees the perceptions by teacher participants (and students when they are considered), that there are positive effects of introducing assessment for learning in classroom practice. Some of the comments from a government report in England illustrate these perceptions, particularly the need for clear information about learning goals and providing challenge.⁵⁰

The opinion of a Physical Education teacher:

'Sharing clear learning objectives supported the progress they [the pupils] made. We differentiate the learning objectives and this provides students with challenges that make sure they push themselves even further. This also encourages them to make their own challenges. As a result of this, plus other work we've done on AfL (notably peer and self-assessment) the higher-ability students are now very confident. They use opportunities to contribute orally in lessons. They will often take a warm-up, coach others and act as referee.'

Comments from pupils:

'If we don't know what we are expected to learn, how do we know if we've got there?'

'When I revise I go straight to the learning objectives and know where to look to revise rather than having to go through loads and loads of pages.'

'They put learning objectives on the board and explain them – I understand more so I can do it. It makes you learn more. I learn what to do before I start... I feel more in control.'

⁴⁷ Webb and Jones (2010: 166).

⁴⁸ Tierney and Charland (2007); Torrance and Pryor (2001); Brookhart (2001).

⁴⁹ Among these can be mentioned Brookhart (2001); Weeden and Winter (1999); MacPhail and Halbert (2010), as well as the reports of DfES (2007) and Kirton et al. (2007).

⁵⁰ DfES (2007: 22).

Most studies of AfL have been implemented as relatively small-scale professional development programmes, with voluntary participants who are keen to try out the new approach. Thus, most studies have worked to a greater or lesser extent with ideal groups in ideal circumstances, which are not always the ones found in the wider population of teachers and schools. An inspection report in England, for example, studied 43 schools and produced evidence that the impact of AfL was outstanding only in 5 schools and good in another 11 schools, while in 27 of them its impact was evaluated as only satisfactory or in some cases inadequate.⁵¹ It was noted that only a few schools of those inspected were actually putting AfL into practice and found that of those that were implementing it, many were not introducing it in accordance with the way advocates suggested was most desirable. That is why the practical feasibility of the approach has become an issue requiring further research in relation to this way of understanding assessment. In most studies the evidence indicates the way various contextual features of school and classroom conditions and also educational system contexts can act as obstacles for promoting and embedding AfL approaches into classroom practice.⁵²

The table following provides a summary of practical suggestions about how schools might make changes and review current practices in order to incorporate AfL approaches into different levels of organisation and day-to-day classroom practice.

⁵¹ Ofsted (2008: 10).

⁵² Tierney and Charland (2007); Thompson and William (2008); Black et al. (2010) and (2004); Azúa and Bick (2009); Gipps (2002); Torrance and Pryor (2001); MacPhail and Halbert (2010); DfES (2007); Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Kirton et al. (2007); Tapan (2001); Gipps et al. (2005); Webb and Jones (2009); McDonald and Boud (2007); Carless (2005); Gioka (2006); Kellard et al. (2008); Stiggins and Arter (2002); Klenowski (2009); Brown et al. (2009); Willis (2008).

A review and audit of current practice should address questions such as:

- What are the existing strengths of the school in promoting student learning?
- How can we find out what these are?
- How can we plan to build on them?
- In what respects are staff close to, or a long way from, the practices that should be encouraged?
- Does this school's existing priorities and plans value students' learning as a top priority?⁵³

Suggestions for subject departments on how to encourage AfL include:

- Have an atmosphere in which teachers are expected to watch others in action – to actively support peer observation.
- Recognise and value current skills and help teachers to identify their current formative practice.
- Have meetings where teachers discuss learning.
- Give teachers time to plan well by encouraging them to mark less, but mark better.⁵⁴

In addition, it is recommended that schools engage in ongoing evaluation of changes in practices linked with AfL approaches. It is recommended that suitable targets could be devised in relation to learning and teaching practices and these would involve observations in classrooms and study of student work.

Examples suggested include:

- At least some questions asked in class should allow ample wait time and lead to involvement of many students in subsequent discussion.
- Homework books should show written comments by teachers together with evidence that students have acted on these comments to improve their work subsequently.
- Discussion by groups of students involving peer assessment of one another's work should be a regular feature in classrooms.⁵⁵

Key idea: Only a relatively few schools have been found to be implementing AfL in accordance with the way advocates suggested is most desirable. The practical feasibility of AfL has consequently become an issue requiring further research.

⁵³ Black et al. (2010: 104-5).

⁵⁴ Black et al. (2010: 94).

⁵⁵ Black et al. (2010: 111-2).

Effects and impact of assessment for learning

The positive effect of the use of assessment to improve student and teacher learning and classroom practice is very rarely questioned in the literature, a fact confirmed by recent reviews.⁵⁶ All the main strategies considered important for AfL (sharing learning goals/criteria, questioning, formative feedback, peer and self-assessment, formative use of summative tests) are evaluated as overwhelmingly positive in terms of their potential to promote improvements in teachers' classroom practice. Any problems identified are attributed to unfavourable contextual conditions such as lack of time,⁵⁷ insufficient application of AfL principles in teaching⁵⁸ and external constraints. The result of this seems to suggest that AfL approaches usually help participating teachers and schools to adopt a more learning-focused stance,⁵⁹ but further research is needed in order to be able to demonstrate the feasibility of AfL in different contexts.

Impact on student achievement

A crucial factor to indicate whether AfL is successful in terms of the claimed benefits of promoting student learning is the extent that student achievement (which is an outcome of learning) can be shown to have increased through the changes derived from the approach. Unfortunately most studies have not examined changes in student attainment (learning gains) directly.

There is consistent evidence that participants in AfL initiatives attribute an increase in students' achievement to the implementation of AfL or of one of the strategies related to it.⁶⁰ For example, one study reports that 78% of participants evaluate the intervention as successful in its impact on primary pupils' attainment⁶¹ and two of the 35 participating schools also reported improvements in summative tests as a result of AfL strategies.⁶² Similarly, another study noted total or partial improvements in student attainment in eight schools involved in an AfL project in terms of secondary students' Key Stage 3 test results, with only one school not being able conclusively to attribute this improvements to AfL.⁶³ However, other sources reviewed indicate that participants, despite their recognition of the positive effect of AfL in different aspects, feel they need more time or better evidence to be able to confidently state that AfL has a positive impact on students' attainment.⁶⁴

It should be recognised that participants engaged in an improvement initiative such as AfL that provides professional development emphasising the benefits to student learning of such assessment approaches, may well find their views influenced by this – which is why other evidence is also desirable.

⁵⁶ Hodgson and Pyle (2010); Tierney and Charland (2007).

⁵⁷ Torrance and Pryor (2001); Condie et al. (2005); Dori (2003); Kirton et al. (2007); Tapan (2001); Webb and Jones (2009); Carless (2005); Gioka (2006); Kellard et al. (2008).

⁵⁸ Hayward and Spencer (2010); Tapan (2001); Webb and Jones (2009); Klenowski (2009).

⁵⁹ MacPhail and Halbert (2010); DfES (2007); Azúa and Bick (2009).

⁶⁰ Ofsted (2008); DfES (2007); Condie et al. (2005); Hayward and Spencer (2010); Kirton et al. (2007); Webb and Jones (2009); Kellard et al. (2008).

⁶¹ Kirton et al. (2007: 612).

⁶² Kirton et al. (2007: 620).

⁶³ DfES (2007: 32).

⁶⁴ Kellard et al. (2008); Condie et al. (2005).

Quantitative evidence not based on participants' perceptions is limited. Some studies have been developed around some specific strategies. An experimental study observed a highly statistically significant difference in students' attainment in external examinations (in four areas examined), and this was attributed to the work of the experimental group with self-assessment strategies.⁶⁵ Other experimental designs have been centred on alternative assessment strategies, but also incorporate some perspectives and strategies of AfL (feedback, student involvement, assessment intertwined with classroom practice, work with higher order thinking skills).⁶⁶ One study identified a statistically significant difference in students' performance, both in low-level assignments (in chemistry and biology) and also more marked significant differences in high-level assignments in both areas.⁶⁷ Another study concluded that a strong positive correlation can be established between AfL strategies and higher test scores.

There is only one quantitative study that has been conducted which was clearly and completely centred on studying the effect of AfL on student outcomes. This produced a significant, but modest, mean effect size of 0.32 in favour of AfL as being responsible for improving students' results in externally mandated examinations.⁶⁸ It must be mentioned, however, that this study has some methodological problems, explicitly recognised by their authors. These are related to the diversity of control groups they considered and the variety of tests included for measuring students' achievement. All this affects the robustness of comparisons within the study.

In short, it can be said there is much evidence that participants perceive positive effects of AfL on student learning and achievement but there have been only a few quantitative experimental studies that test the effects of AfL on attainment outcomes. Thus further research is needed in order to arrive at more definitive conclusions.⁶⁹

Other effects on students

Most literature on the subject refers to claims of the positive effects for students in terms of different areas, such as their role in the classroom, their relationship to learning and their self-concept. In terms of student role, AfL is said to encourage students to become more active and participative in classroom interaction.⁷⁰ As dialogue in a non-threatening environment is promoted, it is suggested that they feel more confident to intervene and propose ideas during the class.⁷¹ Along with this, the practice of peer assessment and self-assessment generally requires students to take a more prominent role in the classroom.

It is further argued that students' relationship to learning changes, in terms of developing a greater awareness of what is expected of them, due to the explicit focus on communicating learning goals and criteria. This shifts their attention from the competition and comparison of themselves with other students to a stronger focus on understanding their own process of learning and their progress according to pre-established learning goals and criteria. Because of this, it is argued that they become more responsible for their own learning progress, which is monitored by them through self-assessment methods.⁷² All this change is held to be beneficial for them in developing further their metacognition skills⁷³ and so enabling them to become more independent learners.

⁶⁵ McDonald and Boud (2007).

⁶⁶ Bartoochi and Keshavarz (2002); Dori (2003).

⁶⁷ Dori (2003: 44).

⁶⁸ William et al. (2004).

⁶⁹ For a detailed critical analysis of quantitative evidence that demonstrates the effectiveness of AfL, see Bennett (2011).

⁷⁰ Torrance and Pryor (2001); MacPhail and Halbert (2010); DFES (2007); Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Willis (2008); Kirton et al. (2007); Tapan (2001); Gipps et al. (2005); Webb and Jones (2009); Bartoochi and Keshavarz (2002); McDonald and Boud (2007); Carless (2005); Kellard et al. (2008); Stiggins and Arter (2002).

⁷¹ Hodgson and Pyle (2010); Torrance and Pryor (2001); Willis (2008); Kirton et al. (2007); Tapan (2001); Webb and Jones (2009).

⁷² Torrance and Pryor (2001); MacPhail and Halbert (2010); DFES (2007); Weeden and Winter (1999); Brookhart (2001); Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Kirton et al. (2007); Tapan (2001); Gipps et al. (2005); Webb and Jones (2009); Bartoochi and Keshavarz (2002); McDonald and Boud (2007); Carless (2005); Kellard et al. (2008); Klenowski (2009).

⁷³ Hodgson and Pyle (2010); Black et al. (2004); Brookhart (2001); Gipps et al. (2005); Bartoochi and Keshavarz (2002); Stiggins and Arter (2002).

In addition, it is suggested that, because AfL is intended to reduce the emphasis given to comparison and competition between students, it is likely that students' motivation will increase, as well as their commitment to achieving their own learning goals. Students are, in this context, mainly concerned about their own progress, so the potentially negative effect of comparison on low-attaining students' self-esteem should diminish, making them feel more confident about their capabilities.⁷⁴

Effects on teachers

There is also agreement in most of the literature about the benefits of AfL for teachers' professionalism and teaching practices. Teachers, like students, change their role in classroom interaction when AfL is introduced.⁷⁵ Their participation is said to shift from the prime concern to be a content deliverer who largely controls the classroom dynamics, to a moderator and facilitator of learning who collaborates with students during the class, supporting and monitoring their progress.

According to an inspection report in England,⁷⁶ these changes in teachers' practice were observed in the minority of schools (5 out of 43) evaluated as outstanding through the following concrete aspects:

- clarity about what and how they wanted students to learn
- careful, but also flexible, planning with objectives based on assessment evidence
- regular revisiting and reinforcement of objectives during lessons
- clear notion of what students could and could not do in order to help them to progress
- good questioning, including moments of the class for drawing learning together
- constructive feedback on students' work.

Key idea: The impact of adopting well understood and well implemented assessment for learning approaches is regarded as wide ranging and essential for good practice, according to the literature surveyed. The problems with the approach are generally attributed to the many difficulties it experiences in the process of becoming well understood and well implemented, a goal which seems to be reached only in a few cases.

⁷⁴ Kirton et al. (2007); Webb and Jones (2009); Carless (2005).

⁷⁵ Torrance and Pryor (2001); MacPhail and Halbert (2010); DFES (2007); Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Willis (2008); Kirton et al. (2007); Tapan (2001); Gipps et al. (2005); Webb and Jones (2009); Bartoochi and Keshavarz (2002); McDonald and Boud (2007); Carless (2005); Kellard et al. (2008); Stiggins and Arter (2002).

⁷⁶ Ofsted (2008: 13).

Guidelines for policymaking and practice

Despite the many claims made for the positive effects of AfL approaches, most of the literature reviewed also highlights major difficulties that are thought likely to arise if this approach is intended to be adopted, particularly if it is to be applied to non-volunteers in a whole system. It is recognised in different studies that the use of AfL by teachers is not frequently observed in practice even when many efforts have been made to spread it at a national level.⁷⁷ This is attributed to the barriers of contextual aspects such as lack of time and competing demands related to accountability in terms of external high stakes tests which have been noted in the research cited earlier. Nonetheless, these problems have not been deeply studied, perhaps because such recognition might put at risk the feasibility of the approach.⁷⁸ Among these problems are:

- lack of commitment from senior staff⁷⁹
- contradictions between the aims of national testing systems and AfL, which can lead to teachers opting to focus their teaching towards the former⁸⁰
- replicability of professional development programmes at a system level and contrast between the ideal conditions of a small-scale programme and a wider dissemination process⁸¹
- lack of appropriate disciplinary knowledge or assessment skills in some teachers⁸²
- superficial understanding of the approach which adversely affects the quality of it in practice⁸³
- resistance of some schools and teachers to promoting greater student participation and less teacher control of the learning process in the classroom.⁸⁴

On the basis of these difficulties and of orientations given in the literature, some guidelines for policymaking and practice are suggested below.

Guidelines for policymaking

1. There should be a careful design of dissemination strategies, considering aspects like feasibility, changes needed in order to facilitate the process, and the stakeholders that must be involved, in order to guarantee an adequate understanding of the AfL approach.⁸⁵
2. Possible contradictions between different policies, which might deter practitioners from engaging with the approach, must be evaluated. Among the most important, the use of national testing and assessment strategies and the role of public examinations must be evaluated in terms of their consistency with AfL, implementing if necessary national assessment systems with a greater degree of teacher involvement.
3. Directly related to the previous ideas, policymakers should give clear information to the educational community and the public about the views on assessment which are to be promoted or prioritised in an education system, in order to avoid confusion and overlapping of approaches among practitioners as well as internal disagreements within schools.

⁷⁷ Black et al. (2010); Thompson and William (2008); Ofsted (2008); Weeden and Winter (1999); Tapan (2001); Gioka (2006); Kellaghan (2004).

⁷⁸ For studies on the implementation of AfL as a policy and the main challenges in this process, see Berry and Adamson (eds) (2011) and Ball et al. (2012).

⁷⁹ Ofsted (2008).

⁸⁰ Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Kirton et al. (2007); Gipps et al. (2005); Black et al. (2010); Black et al. (2004).

⁸¹ Torrance and Pryor (2001); DfES (2007); Dori (2003); McDonald and Boud (2007).

⁸² DfES (2007); Kellaghan (2004); Carless (2005); Gioka (2006); Stiggins and Arter (2002); Ofsted (2008); Thompson and William (2008); Azúa and Bick (2009).

⁸³ Hayward & Spencer (2010); Tapan (2001); Webb & Jones (2009); Klenowski (2009); Ofsted (2008); Azúa & Bick (2009).

⁸⁴ MacPhail and Halbert (2010); DfES (2007); Willis (2008); Kirton et al. (2007); Stiggins and Arter (2002); Thompson and William (2008); Black et al. (2004).

⁸⁵ Gardner et al.'s (2011) diagnosis of AfL as a policy is coincident with this guideline, as it evaluates AfL reforms as generally under-designed.

4. Literature argues that teachers have a need for support in the dissemination process.⁸⁶ This is due to the complexity involved in understanding AfL with the depth required for an adequate transference to practice. In terms of policymaking, some effort must be made to provide support through fostering professional development and internal leaderships in schools.
5. Given contextual diversity, AfL should not be reduced to a rigid model or to the mechanical application of a set of techniques. Enough flexibility should be considered to allow a process of appropriation by practitioners. That is why most of the sources reviewed are in favour of professional development programmes with more active participation of and initiative from teachers,⁸⁷ and give priority to the commitment to and the understanding of the AfL philosophy.⁸⁸
6. If a decision is made to promote AfL in terms of a general policy, monitoring the process of dissemination will be vital to promote quality assurance and generate the necessary conditions within schools. There should also be a commitment to sustain this policy over time and to avoid its replacement by another innovation after a short period, since many school improvement studies have shown that changes take several years to implement and embed.⁸⁹

Guidelines for practice

1. Whole-school commitment to the AfL approach, in terms of introducing a project known and shared by the school community, is seen to be vital to successful uptake.⁹⁰ A group of teachers with no support from parents or senior staff may find it difficult to gain acceptance or understanding for their changes to practice and this is likely to lead them to abandon the new approach, despite their professional development experience. Involvement and leadership of senior staff is especially important to embed AfL, as they are the ones who can promote a reform at a whole-school level, supporting and stimulating teachers in the process of its implementation.⁹¹
2. This commitment at a school level also involves generating the right conditions for an adequate dissemination of AfL, in order to avoid common problems observed in practice by researchers. Among these conditions, there should be:
 - a. a regular space and time for teachers to reflect and discuss with their colleagues
 - b. an explicit shift in the assessment culture of the school
 - c. a clear message about new possibilities in classroom interaction for teachers to know they will not be negatively judged because of innovations
 - d. some agreement on the way in which AfL is intended to interact with the existing marking system of the school and with external assessment procedures.
3. Clarity of messages is also relevant at the school level. A change towards assessment for learning should be informed and explained to all stakeholders, including students and parents. It is recognised in the literature that teachers', students' and parents' beliefs about assessment could generate resistance to change, so exploring and making explicit those beliefs becomes crucial before the actual implementation begins.⁹²

⁸⁶ MacPhail and Halbert (2010); DfES (2007); Condie et al. (2005); Dori (2003); Kirton et al. (2007); Tapan (2001); McDonald and Boud (2007); Carless (2005); Kellard et al. (2008).

⁸⁷ Torrance and Pryor (2001); MacPhail and Halbert (2010); DfES (2007); Condie et al. (2005); Dori (2003); Hayward and Spencer (2010); Kirton et al. (2007); Webb and Jones (2009); Carless (2005); Kellard et al. (2008); Stiggins and Arter (2002).

⁸⁸ Torrance and Pryor (2001); DfES (2007); Condie et al. (2005); Hayward and Spencer (2010); Willis (2008); Gipps et al. (2005); Webb and Jones (2009); Kellard et al. (2008); Stiggins and Arter (2002); Klenowski (2009).

⁸⁹ The need for understanding the implementation of AfL as a complex and slow process is highlighted by MacPhail and Halbert (2010); Hayward and Spencer (2010); Willis (2008) and Webb and Jones (2009), as well as the ARG.

⁹⁰ Ofsted (2008); DfES (2007); Tapan (2001); Kellard et al. (2008).

⁹¹ Ofsted (2008); Kellard et al. (2008).

⁹² Relevance of school community beliefs on assessment is highlighted by Torrance and Pryor (2001); Brown et al. (2009); Brookhart (2001); Willis (2008); Tapan (2001); Carless (2005).

4. Teachers should attend well developed and recognised professional development programmes that help them to develop a flexible and deep understanding of the sense of the approach, and not just to accumulate a set of techniques they might apply without the required level of reflection. Ideally, these programmes should include discussion from a disciplinary perspective. The process of implementation should be monitored in the school, with the purpose of giving appropriate support and ensuring the approach will be adopted in practice and not remain only in the realm of theory.
5. Studying students' perspectives and attitudes and incorporating their evidence into the development and implementation of AfL is necessary if firm evidence of its impact and potential to enhance learning and improve educational outcomes is to be generated.
6. Further research is needed to establish evidence of the effects of AfL on a range of student educational outcomes such as academic self-concept, attainment, motivation and engagement behaviour in class.

In general terms, advocates of AfL argue that there is a need for greater understanding of the complexity of taking AfL approaches into practice. They suggest this involves recognising the process is a slow and gradual one. The kind of understanding and appropriation required for an effective implementation of AfL is unlikely to be realised in a short time period.

Criticism of assessment for learning

Some of the problems with assessment for learning have been already mentioned in this document. These are especially related to the feasibility of the approach. Critics recognise this aspect requires further rigorous research, due to the frequency with which contextual aspects have emerged as obstacles in different studies.⁹³

Further research is needed to identify more clearly the relative merits of different strategies; and a particular focus should be given to measuring the impact on a range of student outcomes including attainment, attitudes, classroom behaviour, engagement and motivation, and self-concept as a learner. In addition, further study of students' perceptions and opinions of this method of assessment is needed.

Additionally, three recent critical reviews⁹⁴ highlight some new aspects to be taken into account. These reviews consider the lack of theoretical precision as a major problem in research processes carried out by the ARG in England.

⁹³ Tierney and Charland (2007).

⁹⁴ Taras (2009); Dunn and Malvenon (2009); Bennett (2011).

In terms of theoretical problems, the reviews raise the following concerns:

1. The ARG's research assumes a great difference between summative and formative assessment: such a difference was not present in the initial literature on formative assessment. This distinction originated in a gradual shift of definition from a focus on assessment *processes* to a focus on assessment *purposes*. From the perspective of processes, summative and formative assessments are aspects of the same procedure; from the perspective of purposes, they become two separate activities with different functions.
2. Along with the previous confusion, in different texts and different places of the same text, the ARG authors define formative assessment in different ways and with different emphasis, which contributes to a lack of clarity about the concept and, also distorts results from research, due to the lack of certainty about what is being precisely considered as the object of study in each case.
3. The interrelation between formative assessment, summative assessment and self-assessment is not explicitly detailed by the ARG researchers. Something similar occurs with connections between theoretical principles and practical aspects.
4. Lack of theoretical precision affects the use of the concept of AfL in practice, generating inconsistencies that would not have appeared if a clear and single definition was used. It also questions some of the research evidence provided by the ARG, as not all of it is consistent with the definitions of formative assessment given by the authors.

The critical reviews suggest that advocates of AfL have inappropriately generalised results from non-representative samples to the whole population, as it is the case with one research project centred on students with special educational needs. Another focused on students in highly vulnerable contexts. Other studies have methodological problems, such as absence of control over possible incidence of external factors in research results, or generalisations made on the basis of small numbers of teachers, to mention just some of them.

Glossary

Colouring squares for goal statements: Self-assessment technique which involves students colouring goal statements according to the level of confidence they feel in achieving that goal.

Comment-only marking: A system of feedback centred on giving students written descriptive comments about their work instead of a mark.

Concept mapping: In the context of AfL asking students to build concept maps can be useful for making misconceptions explicit and thus providing the teacher with the opportunity to give adequate feedback.

Informative and descriptive feedback: Opposed to evaluative feedback which is centred on brief comments based on general judgements about students' work (e.g. 'good', 'well done'), descriptive or informative feedback is understood as being centred on giving qualitative information to students about the well-accomplished aspects of their work, the ones that need improvement, and also suggests ways to improve in future activities.

Jigsawing: Method of collaborative learning that involves one student acting as the expert in some topic and teaching it to his or her peers, who in turn have to teach about the topic in which they have become experts.

Now and next steps / Now and next time tasks: A feedback method which consists of giving students a positive statement about their work along with an expected action for learning improvement to be done 'now' as well as an action to be done in a new task.

Oral and written feedback: Both types of feedback differ in the kind of communication established between teacher and student, oral feedback being more direct and interactive than written feedback. That is why when written feedback is given it is recommended to include moments for reading, understanding and discussing comments in order to make sure students are able to integrate the feedback they receive and apply it in new situations.

Paired marking: Peer-assessment method in which two students interchange their work and assign each other a mark justifying their decisions on the basis of shared criteria.

Plenary self-evaluation: Students discuss the learning gained in a specific lesson in the context of a plenary session, ideally in a non-threatening environment where participation and dialogue predominate.

Question setting: Collaborative learning method that consists of students setting questions for their peers.

Self-assessment journals: The student registers in a journal his or her own progress in learning, making comments on the difficulties, misconceptions and advances he or she makes.

Self-marking: Self-assessment technique in which students assign themselves a mark for their work and are able to justify their decision.

Traffic-lighting: A technique used by students to give account of their understanding about a specific content. It relates a traffic-light colour to a level of understanding. Red means low understanding, yellow indicates some aspects needing further support and green a high level of understanding. This tool has been used in practice for self-assessment and peer assessment as well as for feedback on the quality of teaching.

Two stars and a wish: This tool was developed in some contexts as a practical solution to give students precise and useful feedback. It consists of giving students positive comments about their work plus one suggestion to improve it. It can be used by the teacher or as a tool for peer-assessment.

References and further reading

* References included in the paper are marked with an asterisk.

Adamson, B. (2011) 'Embedding assessment for learning'. In: R. Berry and B. Adamson (eds) *Assessment reform in education: policy and practice*. London: Springer, 197-203.*

Allal, L. and Mottier, L. (2005) 'Formative assessment of learning: a review of publications in France'. In: OCDE, *Formative assessment: improving learning in secondary classrooms*. Centre for Educational Research and Innovation (CERI).*

ARG (2002) *Assessment for learning: 10 principles: research-based principles to guide classroom practice*. Assessment Reform Group. <http://www.aaia.org.uk/content/uploads/2010/06/Assessment-for-Learning-10-principles.pdf>* [Accessed: July 2013]

Azúa, X. and Bick, M. (2009) 'Formación de profesores en evaluación para el aprendizaje' [Teacher education on assessment for learning]. In: C. Sotomayor and H. Walker (eds) *Formación continua de profesores*. Santiago de Chile: Editorial Universitaria. 263-283.*

Ball, S., Maguire, M. and Braun, A. (2012) *How schools do policy: policy enactments in secondary schools*. London: Routledge.*

Barootchi, N. & Keshavarz, M.H. (2002) 'Assessment of achievement through portfolios and teacher-made tests'. *Educational research*, 44(3), 279-288.*

Bell, B. & Cowie, B. (2001) 'A review of the relevant literature'. In: *Formative assessment and science education*. Dordrecht: Kluwer Academic Publishers.*

Bennett, R.E. (2011) 'Formative assessment: a critical review'. *Assessment in education: principles, policy & practice*, 18(1), 5-25.*

Berry, R. and Adamson, B. (eds) (2011) *Assessment reform in education: policy and practice*. London: Springer.*

Black, P. and Wiliam, D. (1998) 'Assessment and classroom learning'. *Assessment in education: principles, policy & practice*, 5(1), 7-74.

Black, P., Harrison, C., Lee, C., Marshall, B., Wiliam, D. (2010) *Assessment for learning, putting it into practice*. London: Open University Press.*

Black, P. with the King's College London Assessment for Learning Group (C. Harrison, C. Lee, B. Marshall, D. Wiliam) (2004) *The nature and value of formative assessment for learning (Draft)*. London: King's College.*

Brookhart, S.M. (2001) 'Successful students' formative and summative uses of assessment information'. *Assessment in education*, 8(2), 153-169.*

Brown, G.T.L., Kennedy, K.J., Fok, P.K., Chan, J.K.S., Yu, W.M. (2009) 'Assessment for student improvement: understanding Hong Kong teachers' conceptions and practices of assessment'. *Assessment in education: principles, policy & practice*, 16(3).*

Butler, R. (1987) 'Task-involving and ego-involving properties of evaluation: effects of different feedback conditions on motivational perceptions, interest, and performance'. *Journal of educational psychology*, 79(4), 474-482.

Carless, D. (2005) 'Prospects for the implementation of assessment for learning'. *Assessment in education: principles, policy & practice*, 12(1), 39-54.*

- Casanova, M.A. (1999) *Manual de evaluación educativa* [Handbook on educational evaluation]. Madrid: La Muralla.
- Condie, R., Livingston, K., Seagraves, L. (2005) *Evaluation of the assessment for learning programme: final report*. Glasgow: Quality in Education Centre, University of Strathclyde.*
- DfES (2007) *Assessment for learning: 8 schools project report*. London: DfES Publications.*
- Dori, Y.J. (2003) 'From nationwide standardized testing to school-based alternative embedded assessment in Israel: students' performance in the Matriculation 2000 project'. *Journal of research in science teaching*, 40(1), 34-52.*
- Dunn, K. and Mulvenon, S. (2009) 'A critical review of research on formative assessment: the limited scientific evidence of the impact of formative assessment in education practical assessment'. *Research & evaluation*, 14(7), 1-11.*
- Flórez Ochoa, R. (1999) *Evaluación pedagógica y cognición* [Pedagogical evaluation and cognition]. Bogotá: McGraw-Hill.
- Gardner, J., Harlen, W., Hayward, L., Stobart, G. (2011) 'Engaging and empowering teachers in innovative assessment practice'. In: R. Berry and B. Adamson (eds) *Assessment reform in education: policy and practice*, London: Springer, 105-119.*
- Gioka, O. (2006) 'Assessment for learning in physics investigations: assessment criteria, questions and feedback in marking'. *Physics education*, 41(4), 342-346.*
- Gipps, C. (2002) *Beyond testing (towards a theory of educational assessment)*, 5th edition. London: Routledge Falmer.*
- Gipps, C., McCallum, B., Hargreaves, E., Pickering, A. (2005) 'From TA to Assessment for Learning: the impact of assessment policy on teachers' assessment practice'. Paper presented at the British Educational Research Association Annual Conference, University of Glamorgan, 14-17 Sept 2005.*
- Hallam, S. et al. (2004) *Assessment is for learning development programme: evaluation of Project 1: support for professional practice in formative assessment: final report* [electronic resource]. London: Institute of Education.
- Hargreaves, E. (2011) 'Teachers' feedback to pupils: "Like so many bottles thrown out to sea"'. In: R. Berry and B. Adamson (eds) *Assessment reform in education: policy and practice*. London: Springer, 121-133.*
- Hattie, J.A. (2005) *What is the nature of evidence that makes a difference to learning?* Research conference proceedings (pp.11-21). Camberwell, VIC: Australian Council for Educational Research. Available at <http://www.acer.edu.au>
- Hayward, L. and Spencer, E. (2010) 'The complexities of change: formative assessment in Scotland'. *Curriculum journal*, 21(2), 161-177.*
- Hodgson, C. and Pyle, K. (2010) *A literature review of assessment for learning in science*. Slough, UK: National Foundation for Educational Research.*
- Kellaghan, T. (2004) *Assessing student learning in Africa*. Washington, DC: World Bank.
- Kellard, K., Costello, M., Godfrey, D., Griffiths, E., Rees, C. (2008) *Evaluation of the developing thinking and assessment for learning development programme*. Welsh Assembly Government.*
- Kirton, A., Hallam, S., Peffers, J., Robertson, P., Stobart, G. (2007) 'Revolution, evolution or a Trojan horse? Piloting assessment for learning in some Scottish primary schools'. *British educational research journal*, 33(4), 605-627.*

- Klenowski, V. (2009) 'Assessment for learning revisited: an Asia-Pacific perspective'. *Assessment in education*, 16(3), 263-268.*
- Ko, J. & Sammons, P. with Bakkum, L. (2013) *Effective teaching: a review of research and evidence*. Reading: CfBT Education Trust.*
- Köller, O. (2005) 'Formative assessment in classrooms: a review of the empirical German literature'. In: OCDE. *Formative assessment: improving learning in secondary classrooms*. Centre for Educational Research and Innovation (CERI).*
- MacPhail, A. and Halbert, J. (2010) '“We had to do intelligent thinking during recent PE”: students' and teachers' experiences of assessment for learning in post-primary physical education'. *Assessment in education: principles, policy & practice*, 17(1), 23-39.*
- McDonald, B. & Boud, D. (2007) 'The impact of self-assessment on achievement: the effects of self-assessment training on performance in external examinations'. *Assessment in education*, 10(2), 209-220.*
- OCDE (2005) *Formative assessment: improving learning in secondary classrooms*. Centre for Educational Research and Innovation (CERI).
- Ofsted (2008) *Assessment for learning: the impact of National Strategy support*. London: Ofsted.*
- Oman Ministry of Education (2001) *National report on the development of education in the Sultanate of Oman*. Muscat: Ministry of Education.
- Rowe, K. (2006) 'Effective teaching practices for students with and without learning difficulties; issues and implications'. *Australian journal of learning disabilities*, 11(3), 99-115.
- Stiggins, R. and Arter, J. (2002) 'Assessment for learning: international perspectives'. Proceedings of an international conference, Chester, UK, 18 September, 2001. Corporate source: Assessment Training Inst. Inc., Portland.*
- Tam, H.P. and Lu, Y.J. (2011) 'Developing assessment for learning in a large-scale programme'. In: R. Berry and B. Adamson (eds) *Assessment reform in education: policy and practice*. London: Springer, 185-196.*
- Tan, K. (2011) 'Assessment for learning reform in Singapore – quality, sustainable or threshold?' In: R. Berry and B. Adamson (eds) *Assessment reform in education: policy and practice*. London: Springer, 75-87.*
- Tapan, S. (2001) *A study of present learning assessment practices to ensure quality of primary education. Basic education studies, 2000-2001*. Dhaka: UNESCO Office Dhaka.*
- Taras, M. (2009) 'Summative assessment: the missing link for formative assessment'. *Journal of further and higher education*, 33(1), 57-69.*
- Thompson, M. and Wiliam, D. (2008) 'Tight but loose: a conceptual framework for scaling up school reforms'. In: C. Wylie (ed) *Tight but loose: scaling up teacher professional development in diverse contexts*. Princeton, NJ: Educational Testing Service (ETS).*
- Torrance, H. and Pryor, J. (2001) 'Developing formative assessment in the classroom: using action research to explore and modify theory'. *British educational research journal*, 27(5), 615-631.*
- Van der Werf (2006) 'General and differential effects of constructivist teaching'. Keynote presentation at the annual meeting of the International Congress for School Effectiveness and Improvement, Florida, 3-6 January 2006.

Webb, M. and Jones, J. (2009) 'Exploring tensions in developing assessment for learning'. *Assessment in education*, 16(2), 165-184.*

Weeden, P. and Winter, J. with Broadfoot, P. et al. (1999) *Learners' expectations of assessment for learning nationally: report for the Qualifications and Curriculum Authority*. Bristol: Graduate School of Education, University of Bristol.*

William, D., Lee, C., Harrison, C., Black, P. (2004) 'Teachers developing assessment for learning: impact on student achievement'. *Assessment in education: principles, policy & practice*, 11(1), 49-65.*

William, D. (2009) 'Assessment for learning: why, what and how?' Inaugural professorial lecture, London: Institute of Education, University of London.

Willis, J. (2007) 'Assessment for learning – why the theory needs the practice'. *International journal of pedagogies and learning*, 3(2), 52-59.*

Willis, J. (2008) 'Assessment for learning: a socio-cultural approach'. In: AARE 2008 International education research conference: Brisbane: papers collection: Conference of the Australian Association for Research in Education, 30 November – 4 December 2008, compiled by P.L. Jeffrey. Melbourne: Australian Association for Research in Education.





CfBT Education Trust
60 Queens Road
Reading
Berkshire
RG1 4BS
0118 902 1000
www.cfbt.com