

Summative assessment: the missing link for formative assessment

Maddalena Taras*

University of Sunderland, Sunderland, United Kingdom

Assessment for learning is increasingly part of accepted orthodoxy, with massive government funding in England, is central to national assessment in Wales, and an export to the USA. Black et al.'s *Assessment for learning: Putting it into practice* (2003), the 'bible' of assessment for learning, is set reading for trainee teachers across the UK, and this text is increasingly a staple diet for all interested in assessment for learning. As such it has an important impact on all involved in the teaching and learning process. Despite this, there has been little discussion of either the paradigm or the definitions which inform it. This article examines the definitions of formative assessment and the theoretical premises of assessment for learning exemplified here and how they impact on the practices described. It finds a lack of alignment and coherence in the rationale of the theory, and contradictions which ensue in the practice. One solution is a paradigm shift basing definitions of formative and summative assessment on processes of assessment and not on functions. Functions remain as a basic epistemological premise of assessment.

Keywords: assessment; summative; formative

Introduction

This article examines practices and theoretical premises of assessment for learning in Black et al. (2003). This book is largely a distillation of the work of Black and Wiliam, which lays the foundation for theory (Wiliam and Black 1996; Black and Wiliam 1998a; Wiliam 2000) and practice (Black and Wiliam 1998b; Black 2003c), and as such it is representative of the assessment for learning paradigm. The theory and practice in the earlier work has been evaluated by Taras (2007c), and this critique builds on that.

It is perhaps necessary to put this critique into perspective. The work of Black, Wiliam and the Assessment for Reform group, of which Black et al. (2003) will serve as an example, represents important developments in theory and practice. Furthermore, much work at the chalk face to disseminate and roll out evidence-based practice involves complex scenarios with multiple protagonists. This article has a much narrower remit. Nonetheless, since sound theory addresses the very heart of assessment procedures, even minor changes could have considerable and far-reaching repercussions on quality and efficiency in practice.

Though this paper is epistemologically critical of the theoretical framework of Black and Wiliam's work, exemplified here by Black et al. (2003), it is necessary to keep in mind that they have endeavoured to provide a practical and ethical focus to supporting learners and learning.

Scriven's (1967) distinction of formative assessment (FA) and summative assessment (SA), Sadler's (1989) theory of formative assessment, and Taras' (2005)

*Email: maddalena.taras@sunderland.ac.uk

discussion of assessment terminology support this article. The principles of their work that are directly relevant will be discussed briefly for the benefit of non-specialist readers.

A critical distinction which is made in this paper and which is seen as emblematic of the problems arising in the assessment for learning paradigm is that between functions and processes of assessment. Scriven (1967) and Sadler (1989) are used to clarify this distinction.

The process of assessment

The process of assessment is the mechanism which carries out a judgement. A judgement cannot be made within a vacuum, and therefore points of comparison (i.e. criteria and/or standards) are necessary and in constant interplay. With implicit parameters, it may be difficult for others analysing the assessment to understand the salient points being prioritised. Explicit parameters go some way towards creating a shared forum for assessment and therefore facilitating transparency of process, although within any given context, meanings can and do vary between individuals (for a discussion of criteria, see Sadler 1989, 2005).

Scriven (1967), in the context of course development, made the initial distinction between SA and FA. He notes that the process of assessment is a single process which makes a judgement according to criteria and standards; this is SA, and is always the first part of any assessment process. FA is an additional step which follows SA and necessitates feedback indicating the possible ‘gap’ in addressing the criteria or the required standard (Sadler 1989); finally, the learner must use the information in future activities (Ramaprasad 1983).

Therefore, in addition to having functions and products of assessment, both SA and FA are processes. It follows that an assessment can be uniquely summative when the assessment stops at the judgement. It is not possible for assessment to be uniquely formative: SA may be implicit and only the formative focus made explicit,

Table 1. Comparative steps for summative and formative assessment.

Scriven 1967	Sadler 1989
1. a weighted set of goal scales	1. concept of standards, goals or reference levels
2. gathering and combining of performance data	
3. to yield either comparative or numerical ratings	2. compare actual level with standard
4. in the justification of	
(a) the data-gathering instruments	
(b) the weightings	
(c) the selection of goals	
1+2+3+4=SA	1+2=SA
	Feedback possible only after SA
	3. appropriate action to close the gap
	1+2+3=FA

Source: Taras 2005, 473.

Downloaded At: 23:32 17 December 2010

or both can be explicit. Black and Wiliam (1998a, 8) favour implicit, while Sadler (1998, 2, 53) requires explicit expression of SA; however, Sadler lets this be inferred rather than stating it explicitly (Taras 2005).

The functions of assessment

The functions of assessment refer to its intended use: a use which cannot be guaranteed (Wiliam 2000). The choice of function does not impinge on the actual process of assessment, but it will affect the choices of the parameters of the assessment. Therefore, the criteria and, to a lesser extent, goals and standards will be influenced by the choice of function. Functions are many and can be combined into multiple uses (Black and Wiliam 1998a). Often social functions predominate over educational ones (Broadfoot 2002); for example, creating hierarchies of selection has often been prioritised over assessment which represents learners' levels of expertise. Society rightly makes judgements; the fear of misuse of these judgements has distorted our view of assessment (Scriven 1967, Rowntree 1987).

Since Scriven drew the distinction between SA and FA, a gradual separation of the two into mutually exclusive entities based on the differences in functions of assessment has evolved in the literature. Bloom, Hastings, and Madaus (1971) seem to be the first to have created this dichotomy (Wiliam and Black 1996, 537). Bloom, Hastings, and Madaus (1971, 54) 'borrowed' the term 'formative' from Scriven and used it to mean the provision of feedback on tests from small learning units which make up a mastery learning framework. However, and more importantly, they introduce new parameters to differentiate between SA and FA: 'The distinguishing characteristics have to do with purpose (expected uses), portion of course covered (time), and level of generalization sought by the items in the examination used to collect data for the evaluation' (Bloom, Hastings, and Madaus 1971, 61).

Although not absolute, they note that the last is perhaps the feature which differentiates the two most sharply (Bloom, Hastings, and Madaus 1971, 62). This begs the question as to what constitutes level of generalisation and how it can be determined. Torrance (1993, 335–6) has criticised this model as behaviouristic. Also, importantly, this does not correspond to the distinction which was first made by Scriven.

Black and Wiliam have, both as a team and separately, worked extensively to promote FA and support learning over the pressures and constraints of SA for validation and certification. They have used the dichotomy of SA and FA as created by Bloom, Hastings, and Madaus (1971) in their definitions of FA (Black and Wiliam 1998a, 8; Black et al. 2003, 2), as will be seen in the discussion to follow. They also use the model of formative assessment and feedback developed by Sadler (1989), which follows the model originally proposed by Scriven (Taras 2005).

Many of the problems highlighted in this article emanate from this focus on functions as adopted from Bloom, Hastings, and Madaus (1971).

This article claims that the eclipse of the clarity of processes of assessment and their replacement with a focus on the *functions* of assessment (however well intentioned socially and politically [Taras 2005, 2007a]) has resulted in many of the problems that are highlighted in this article. A return to basing assessment discussions on processes will go a long way to remedying these problems.

Background

Critiques of the work of Black and Wiliam

There are three main critiques of the work of Black and Wiliam to date: they have focused on the pedagogic environment (Perrenoud 1998), replicability (Smith and Gorard 2005b) and the theoretical foundations (Taras 2007c).

Perrenoud (1998) criticises the lack of theoretical pedagogic context when Black and Wiliam are discussing the successes of assessment for learning interventions. This would seem to accord with a tendency noted by Tight (2004), who found a dearth of theoretical focus in a trawl of educational research papers:

there are strong pressures on those working in educational departments to focus their research energies on identifying what works best in the classroom or lecture theatre ... the demand for evidence-based practice gives relatively little priority to theory. (Tight 2004, 406)

Black and Wiliam (2006) respond to Perrenoud and provide an explanation for the success of their classroom interventions to improve learning in schools. The chapter does not, however, provide 'a theory of formative assessment' as the title would seem to indicate.

Smith and Gorard (2005a) report a study which purports to replicate one of the four procedures of formative assessment – that is, 'feedback without a mark' to support learning. They note, first, the difficulty of replication, and second, that the results were the opposite to those reported in the work of Black and Wiliam. A discussion across three issues of *Research Intelligence* (May, August and November 2005) brought to light the fact that Smith and Gorard had misunderstood the principles and ethos of this formative assessment procedure, and that they had not communicated its reality to either staff or pupils involved in the study, all of which Smith and Gorard acknowledge (2005b). This would seem to suspend the criticism that interventions from assessment for learning are difficult to roll out and duplicate. Furthermore, the massive body of empirical work which has been carried out across the world would serve to indicate the reality of being able to both roll out and duplicate aspects of the work.

The third critique by Taras (2007c) examines the early work of Black and Wiliam which serves as a basis for the theoretical development of the assessment for learning paradigm, and it highlights anomalies and contradictions. The inconsistencies highlighted relate to the definitions of summative and formative assessment on the one hand, and to the definitions and examples of 'processes, functions and systems' on the other.

This article critically evaluates Black et al. (2003), which has become a standard manual for researchers, lecturers, teacher-trainers, teachers, and student-teachers. This 'bible' of formative assessment both presents an overview of the theory and also links the theory to the practical interventions and the empirical studies which have ensued. Two aspects of this book would confirm the necessity for a critique (particularly after the anomalies in theory signalled by Taras 2007c) of, first, the scale of the dissemination and influence of the ideas in it, and, second, its use for models of good practice. This article engages primarily with theory because it serves to evaluate and explain the cogency of practice, and, as a consequence, open the means to improve both practice and theory.

A critique of Black et al. (2003)

Practice always takes place within a theoretical framework, whether this is expressed implicitly or explicitly. This can best be seen by analysing the definitions of FA used. Black et al. (2003, 121) state: 'The clarification of the concept of formative assessment has been one of the significant features of our work.'

By examining both the definitions and the concepts of formative assessment in this book, it is difficult to agree that this has taken place. Two different processes of assessment are presented which do not integrate or support each other. The first process is in the context of classroom interaction and is focused on the teacher as the instigator and controller of the process. This definition of FA is consistent with that of Black and Wiliam (1998a, 8): 'all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities'.

Thus in both cases FA serves to change classroom learning and teaching processes. Although Scriven (1967) notes that the assessment process can be applied to everything and anything, here FA seems to be limited to classroom interaction. Black and Wiliam seem to provide an equal or balanced responsibility in FA for both teachers and students, whereas Black et al. weight the responsibility firmly onto the teacher: 'It has to be within the control of the individual teacher and, for this reason, change in formative assessment practice is an integral and intimate part of a teacher's daily work' (Black et al. 2003, 2).

From this it follows that FA is part of teaching methodology and has more to do with teachers than with learners, thus seeming to contradict the initial definitions.

The second process of FA in the book, which seems to be a dramatic conceptual leap, does not correspond to the initial definition. They discuss Sadler's (1989) central concept of feedback (Black et al. 2003, 13–15). This states that information on some attribute is compared with the desired level to identify a gap; this information is used to alter the gap. 'With small changes of terminology, the above four steps could be a description of formative assessment' (Black et al. 2003, 15).

This is Sadler's definition of FA. This differs from the initial definitions by Black et al. (2003) and by Black and Wiliam (1998a) in that, first, it relates to learners (i.e. it assesses their work and not teachers' classroom processes), and second, it is about the process of assessment (i.e. four steps) which relates to product and is not about teacher methodology in the classroom (Wiliam 2000, 15; Black 2003c, 2; Black et al. 2003, 15, 121).

Sadler's article describes and discusses formative assessment, of which formative feedback is an integral part and which takes place when the assessment has been completed; the assessment can be carried out by the tutor or the learner, but the feedback must be used by the learner. It follows that, to understand the feedback, learners must also understand the difference between the initial assessment and what is required; therefore, learners also need to be part of the assessment process. This initial assessment is SA. By not naming the initial step of SA, the whole process of FA seems to have become confused by Black et al. (2003).

Subsequently, they note that their work has clarified the concept of FA (Black et al. 2003, 122); this again seems to return to the original definitions cited at the beginning of the section, as they do intermittently throughout the book.

In a book which presents both a theoretical and a practical framework of FA, it is surprising that, first, two separate and disparate definitions are presented, second,

these are not acknowledged as such, and third, the interrelationship between the theory and the practice is not discussed, particularly when the different definitions of FA subsume differing responsibilities and roles for tutors and learners.

Research to support the FA framework for the study

Black et al. (2003, 2), report on 48 teachers in six schools, on a two-and-a-half-year research project. They use Black and Wiliam's (1998a) research evidence to support the use of FA: four indicative examples provide evidence of improvements due to FA. In principle this is a useful exercise. In practice, due to the lack of clarity in three areas – notably, the definition of FA, the process of FA, and the role of the protagonists – it is difficult to reconcile these four examples with the framework for FA. Importantly, these four summary reports represent over 40 research studies.

The first example is of self-assessment as FA. The students in the research quoted were involved in choosing 'learning tasks' and 'in assessing their own learning outcomes' (Black et al. 2003, 7). There are two problems with this choice. First, the students were not actually using self-assessment as FA according to Black et al. and Black and Wiliam's definition. Their definition of FA denotes it as a classroom pedagogic process, whereas self-assessment is represented as an individual process of examining and evaluating product. Second, according to Sadler's (1989) definition of FA, self-assessment is the student equivalent of SA, not FA (Taras 2005). It only becomes FA if students provide themselves with feedback which they subsequently use.

The second example is from the work of Fuchs and Fuchs (1986), who focus on children with mild disabilities. This research looks at the use of feedback from learners to teachers and by teachers to learners. Teachers worked with systematic procedures to review the assessments and take action accordingly. Black et al. (2003) note that there are two central features to report on this. Firstly, 'the striking success of the interactive (i.e. formative), approach, and secondly, that the main learning gains from the formative work were only achieved when teachers were constrained to use the data in systematic ways, ways which were new to them' (Black et al. 2003, 7–8). This is an example of FA according to the definitions at the beginning of this section as proposed by Black and Wiliam (1998a) and Black et al. (2003), and therefore shows internal consistency and coherence. Importantly, by focusing on the process and making it explicit, their definition could be conflated and overlapped with Sadler's to produce a cogent theoretical framework which is all-encompassing.

The third example of research (Bergan et al. 1991) was with five-year-old children. Teachers were trained to use assessment feedback as FA for learners. This uses Sadler's definition of FA again, not FA as classroom process as defined initially by Black et al. (2003).

The fourth example was work by White and Frederiksen (1997) and involves 'discussion to promote reflective assessment, with both peer assessment of presentations to the class and self-assessment' (Black et al. 2003, 8).

It was found that a greater understanding of the assessment process produced higher scores (Black et al. 2003, 9). As with the first example, neither peer nor self-assessment per se can in itself be technically called FA. Again, this is SA according to Sadler's definition of FA, not FA as originally defined by Black et al. In three out of the four examples provided, Black et al. provide evidence which supports Sadler's

definition of FA, not the FA which they describe as being the classroom learning and teaching process.

If FA is teacher led and relates to classroom processes, it is difficult to see why there is a constant switching to Sadler's definition. It would be both logical and efficient to recognise that, first, assessment as a process can and does apply to all contexts and circumstances (Scriven 1967; Rowntree 1987; Taras 2005), which would reconcile the inconsistencies that have manifested themselves in Black et al.; Sadler's definition could be extended and adapted to include the use of the feedback cycle for classroom interaction.

What Black et al. seem to be doing is to have their own definition of FA which is based on the distinction of functions (taken from Bloom, Hastings, and Madaus, 1971) which they relate to classroom interaction and which is under the control of the teacher, but also, seeing the necessity to address FA as learner product (i.e. of their assignments or work), they adopt Sadler's definition. FA can serve all contexts if, first, functions are seen as a separate issue and if Sadler's definition is incorporated into a complete framework of assessment; second, recognising the relationship in the process between SA and FA would reconcile the contradictions which appear when qualifying an activity or process as formative. In practice, much of the feedback may indeed be used, but unless our processes are made clear and explicit, then learning becomes ad hoc, which is not the aim of educational assessment or formative assessment (Sadler 1989).

Theory in practice: formative assessment activities

The teachers involved in Black et al.'s study were given a choice as to which formative activities they wished to integrate into their teaching. Providing a choice is considered, and is important, because teachers must feel that the activities they use will fit their own model of teaching and learning, and harmonise with their personality. This permits them to take ownership of the new processes: 'with support, teachers can transform research results into new and effective practices' (Black et al. 2003, 56).

Initially, four ideas were taken from 'Inside the Black Box' (Black and Wiliam 1998b) – that is to say, questioning, feedback, sharing criteria, and self-assessment (Black et al. 2003, 30). As the project developed, criteria were subsumed under the feedback and self-assessment areas, and formative use of summative test was added, principally because teachers found it unrealistic to separate SA and FA. The four final areas which were developed during the project were questioning, feedback through marking, peer and self-assessment, and the formative use of summative tests (Black et al. 2003, 30–57). These will be analysed briefly and examined for internal coherence.

Questioning

The basis of change was taken, first, from a study by Rowe (Rowe and Hill 1996) – who found that tutors waited less than a second for a reply before they intervened – and, second, from teachers realising that they tended to ask simple, closed questions 'where recall rather than thinking provided the answer' (Black et al. 2003, 32). By the end of the study, teachers had significantly changed their procedures, attitudes and reasons for asking questions. 'Put simply, the only point of asking questions is to

raise issues about which the teacher needs information or about which the students need to think' (Black et al. 2003, 42).

Teachers found that students became more active participants and that they understood that learning was related to their being able to express and discuss their ideas (Black et al. 2003, 42). The change in the questioning framework seems to have moved learning from a behaviouristic theory where factual recall was prioritised, to a social constructivist perspective taken from cognitive psychology, where a complex framework of factors within a given context permits learners to explore their own understanding. This would also seem to correspond to 'deep' learning as opposed to 'surface' learning (Atkins, Beattie, and Dockrell 1993, 50–7; Black 2001, 76) – metaphors which have become a value judgement on the quality of learning. Due to the interaction and reaction to ideas, the feedback from others during discussion means that learners can adjust their own understanding and opinions in an iterative cycle of learning. This process seems truly what can be called formative feedback which contributes to learning within classroom interaction. In this context it is difficult to break down how SA and FA fit into the equation, since it is the continual iterative absorption and adaptation of personal concepts and opinions which depend on input. However, we can extrapolate that each individual makes an SA of what they hear from other students before their own ideas are adapted and integrated into their own personal framework (i.e. become FA).

Feedback through marking

The second area which supports student learning is based on 'feedback by comments but without marks' (Black et al. 2003, 42–9). It is based on research by Butler (1988), who studied the impact of types of feedback; one supplied marks, one supplied comments, and one a combination of both. Learning gains are best with comments only, as the presence of marks seems to interfere with effective take-up of feedback. What is perhaps surprising from the reaction of the teachers in Black et al.'s study is that many felt they could not engage with this process. This process seemed to have been perceived as an 'all-or-nothing' situation by the teachers and it is surprising that other choices were not found.

Peer and self-assessment by students

As noted in the previous section, technically neither peer nor self-assessment corresponds to FA, but rather to SA. However, the work that teachers carried out will be briefly examined to see if the final feedback and use of this feedback which is necessary for them to become FA have been included in the practice.

The work of Sadler (1989) was used as a starting point for this, although his work principally looks at FA and not at peer or self-assessment (which he classifies as SA). Teachers use student work, that is, their assessments and therefore the product of the process of assessment for the skill of peer and self-assessment. The principle used is that learners cannot achieve goals unless, first, they understand the goals, and second, they can assess what needs to be done to achieve these (Black et al. 2003, 49). Initially in the study, attempts at both were found to be unsuccessful since students lacked skills to judge problems and set targets. Feedback by comments only (see above) on students' work was used as a means of providing students with these skills. Peer assessment was found to be an important complement and even a

prior requirement to self-assessment (this conforms to findings of other research; see Black and Wiliam 1998a). 'Students learn by taking the roles of teachers and examiners of others' (Black et al. 2003, 42).

Interestingly and importantly, peer and self-assessment 'make unique contributions to student learning ... which cannot be achieved in any other way' (Black et al. 2003, 53). From the reporting it is clear that both peer and self-assessment are not used just as SA, but use Sadler's feedback model. This, again, produces FA according to the Sadler model.

The formative use of summative tests

At the start of the project, teachers were advised to separate SA and FA, but teachers saw the two as inextricably linked and therefore refused this separation (Black et al. 2003, 53). Instead, teachers found ways of reconciling the two despite the difficulty of negotiating in-house assessment with externally controlled tests (Black et al. 2003, 55–6). This would support the link between SA, FA and self-assessment provided by Scriven (1967) and Sadler (1989) (Taras 2005).

Three innovations contributed to the formative use of summative tests. First, students structured their reviewing and revision to focus on areas of weaknesses which they subsequently discussed with their peers (Black et al. 2003, 53). Second, students prepared for exams by generating, answering and marking their own questions. The latter follows research by King (1992b) and Foos, Mora, and Tkacz (1994), who note that this improves students' performance. This helped students understand the assessment process and focus their efforts on improving. Third, the innovation was to 'use the aftermath of tests as an opportunity for formative work' (Black et al. 2003, 55); more specifically, this involved teachers focusing on the weak areas of tests and thus focusing on learning. All three innovations involved students understanding and engaging with criteria, understanding the assessment process, and reflecting on and judging their work and that of their peers effectively. Also, and importantly, there is the recommendation for students to re-work exam questions in class (Black et al. 2003, 55). The last step closes the loop and permits FA because students have the possibility of using their learning (Sadler 1989). 'The overall message is that summative tests should be, and should be seen to be, a positive part of the learning process' (Black et al. 2003, 56).

Therefore, to recap, the four final areas of FA which were developed during the project were questioning, feedback through marking, peer and self-assessment, and the formative use of summative tests (Black et al. 2003, 30–57). All but the second clearly show that these areas are indeed FA and that they support learning. The second area – that is, feedback through marking – could easily do so too. However, it is not clear from the book that the final step (i.e. using the feedback) is actually ensured.

Therefore, this article shows that to ensure that FA has taken place, it is necessary to have a clear framework which, first, agrees on the definition of FA, and second, shows the relationship of FA to SA explicitly.

Solving problems with current practice

The anomalies highlighted in Black et al (2003) are: dual definitions of FA; lack of explicit consideration of SA; and lack of explicit links between SA, FA and

self-assessment in theory and practice. The practice suffers from inconsistencies which could have been avoided by a clear theoretical framework. Furthermore, the whole assessment process is never clarified or made explicit (although it does appear in Wiliam and Black 1996 – see Taras 2007a for a full discussion). Replacing functions with processes of assessment provides a new paradigm which makes explicit and transparent the various steps within formative, summative and self-assessment processes (Taras 2005, 2007a,c). This helps to ensure that tutors and learners are clear with what processes have taken place and with how they were accomplished, and so evaluate their effectiveness.

The issues discussed are pertinent to all sectors and all levels of educational assessment. However, an additional problem arises for the post-compulsory sector – both higher education (HE) and the post-compulsory aspects of further education (FE): transfer of aspects of the assessment for learning paradigm exemplified in Black et al. (2003) are increasing, and not just in education departments. In addition to the problems arising because of the inconsistencies signalled, others occur because HE has very different and disparate assessment traditions and processes from those of the compulsory sector (Taras 2007b). Black et al. (2003, 53) signal that teachers refused to separate SA from FA because they saw the two as intrinsically linked. However, a large body of literature argues extensively and passionately for a separation of FA and SA and demonises the latter (Broadfoot and Black 2004), and this is increasingly the dominant discourse and belief. Despite the laudable socio-political reasons for this discourse, it creates a false dichotomy and separation. It has some rationality in the compulsory sector where classroom learning (often called FA) and external exams (SA) are separate (Biggs 1998; Taras 2005).

However, in HE, traditionally, most feedback (and consequently FA) is normally obtained from SA and graded work. Non-graded work which aids learning has also been an optional but not a necessary component of assessment practice. By adopting the demonisation of SA as the ‘Frankenstein monster’ (Broadfoot 2002), HE risks losing the most powerful and central learning support tool it has.

This said, Black et al. (2003) report a positive outcome to their project. A number of reasons can be given for this. Wittgenstein’s idea of salience is relevant here, as is the Hawthorne effect. Discussing criteria of both process and product enlightens learners and initiates them into the tutor’s assessment culture (Sadler 1989); Torrance and Pryor (2001, 623) state that an understanding of task and quality criteria is central to their FA model, and this is also true for Black et al. (2003, 30). Learners are also required to reflect and feedback to the teacher on both the process and the product of their work, and therefore this provides them with practice at making SA judgements and becoming part of the community’s assessment culture (Sadler 1989). Explicit focus on assessment has given students much food for thought, which has had a positive effect on their learning and their work. Missing from some examples of FA research and processes suggested to teachers are explicit points where FA is used by the learners and then re-examined by the teacher.

Conclusion

Black and Wiliam have been protagonists in creating two metaphors and two new concepts: that of getting inside the ‘black box’ and that of ‘assessment for learning’. These have fired the imagination of educationalists from policy to classroom level,

a rare feat. However, the black box is not limited to the classroom; there is a danger that the assessment for learning paradigm may become the new black box. For their work to fulfil its promise, it must also be open to scrutiny and discussion, and this article has begun to prise open the lid so that assessment for learning does not become the Pandora's box of assessment.

Because assessment for learning has become international property, it is an international responsibility to ensure that this new love affair (Black et al. 2003, 123) becomes meaningful and durable. Many questions remain to be answered in future research, of which the following are an indication: How do the two definitions of FA fit into a coherent theory of assessment? How do functions affect assessment processes? How does separating SA and FA help learners and tutors, or not?

This article argues that two disparate definitions of FA are confusing and theoretically untenable unless this separation can be justified both theoretically and in practice.

Also, that it has been difficult to evaluate the coherence of FA practice since the missing link – that is to say, SA – has been left implicit and excluded from the analysis of FA in practice. By making the process of assessment, including SA, explicit, it is possible to re-evaluate the assessment for learning processes in the FA model developed by Black et al. (2003). By focusing on assessment processes as opposed to the functions of assessment, it is on the one hand easier to be explicit in our procedures and, consequently, to improve them, and on the other, to align theories of formative assessment more expediently with practice.

A second important and related issue is that of opening theoretical discussions. Debate is more likely to occur if we remember that discourse is framed as a war zone, so that we should not set off landmines in the haste of our arguments (Taras 2007a).

The third issue is that of engaging explicitly with theory. Without explicit theory, it is difficult to ascertain just how clear or efficient practice is. Tight notes: 'Higher education researchers, for the most part, do not appear to feel the need to make their theoretical perspectives explicit, or to engage in a broader sense in theoretical debate' (Tight 2004, 409).

Theory cannot be neglected if we are to have a mature community of practice (Tight 2004, 409) which provides research in higher education with more credibility and respect.

Notes on contributor

Maddalena Taras' research has focused on a range of assessment issues: self-assessment – developing an original framework and examining issues of power; institutional discrepancies and contradictions in assessment practices and discourses; constraints from language, culture and power impacting on assessment perceptions and practices; and critiquing the 'Assessment for Learning' paradigm, particularly the theoretical framework.

References

- Atkins, M.J., J. Beattie, and W.B. Dockrell. 1993. *Assessment issues in higher education*. Sheffield: Employment Department.
- Bergan, J.R., I.E. Sladeczek, R.D. Schwarz, and A.N. Smith. 1991. Effects of a measurement and planning system on kindergartners' cognitive development and educational programming. *American Educational Research Journal* 28: 683–714.

- Biggs, J. 1998. Assessment and Classroom Learning: a role for summative assessment? *Assessment in Education: Principles, Policy & Practice* 5, no. 1: 103–110.
- Black, P. 2001. Dreams, strategies and systems: Portraits of assessment past, present and future. *Assessment in Education: Principles, Policy & Practice* 8, no. 1: 65–85.
- Black, P. 2003. *Formative and summative assessment: Can they serve learning together?*. In collaboration with the King's College London Assessment for Learning Group, C. Harrison, C. Lee, B. Marshall, and D. Williams. Paper presented at American Educational Research Association (AERA), 23 April, in Chicago, IL. SIG Classroom Assessment Meeting 52.028. <http://www.kcl.ac.uk//depsta/education/hpages/pblackpubs.html>.
- Black, P., C. Harrison, C. Lee, B. Marshall, and D. Williams. 2003. *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.
- Black, P., and D. Wiliam. 1998a. Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice* 5, no. 1: 7–74.
- Black, P., and D. Wiliam. 1998b. Inside the black box: Raising standards through classroom assessment. *PHI Delta Kappan* 80, no. 2: 139–44.
- Black, P., and D. Wiliam. 2006. Developing a theory of formative assessment. In *Assessment and learning*, ed. J. Gardner. London: Sage.
- Bloom, B.S., J.T. Hastings, and G.F. Madaus, eds. 1971. *Handbook on the formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Broadfoot, P. 2002. Editorial: Beware the consequences of assessment! *Assessment in Education: Principles, Policy and Practice* 9, no. 3: 285–8.
- Broadfoot, P., and P. Black. 2004. Redefining assessment? The first ten years of *Assessment in Education*. *Assessment in Education* 11, no. 1: 7–27.
- 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.
- Foos, P.W., J.J. Mora, and S. Tkacz. 1994. Student study techniques and the generation effect. *Journal of Educational Psychology* 86: 567–76.
- Frederiksen, J.R., and B.J. White. 1997. Reflective assessment of students' research within an inquiry-based middle school science curriculum. Paper presented at the annual meeting of the American Educational Research Association (AERA), 24–28th March, in Chicago, IL.
- Fuchs, L.S., and D. Fuchs. 1986. Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children* 53: 199–208.
- King, A. 1992a. Comparison of self-questioning, summarizing, and note-taking review as strategies for learning from lectures. *American Educational Research Journal* 29: 303–23.
- King, A. 1992b. Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist* 27: 111–26.
- Perrenoud, P. 1998. From formative evaluation to a controlled regulation of learning. *Assessment in Education: Principles, Policy and Practice* 5, no. 1: 85–103.
- Ramaprasad, A. 1983. On the definition of feedback. *Behavioural Science* 28: 4–13.
- Rowe, K.J., and P.W. Hill. 1996. Assessing, recording and reporting students' educational progress: The case for subject profiles. *Assessment in Education: Principles, Policy & Practice* 3: 309–51.
- Rowntree, D. 1987. *Assessing students: How shall we know them?* London: Harper and Row.
- Sadler, D.R. 1989. Formative assessment and the design of instructional systems. *Instructional Science* 18: 145–65.
- Sadler, D.R. 1998. Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice* 5, no. 1: 77–85.
- Sadler, D.R. 2005. Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education* 30, no. 2: 175–94.

- Scriven, M. 1967. The methodology of evaluation. In *Perspectives on curriculum evaluation*, ed. R. Tyler, R. Gagne, and M. Scriven. AERA monograph series – curriculum evaluation Chicago: Rand McNally.
- Smith, E., and S. Gorard. 2005a. ‘They don’t give us our marks’: The impact of formative assessment techniques in the classroom. *Assessment in Education: Principles, Policy & Practice* 12, no. 1: 21–38.
- Smith, E., and S. Gorard. 2005b. The dissemination of formative assessment: A short response to Black, Harrison, Hodgen, Marshall and Wiliam. *Research Intelligence* 93: 6.
- Taras, M. 2005. Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies* 53, no. 4: 466–78.
- Taras, M. 2007a. Machinations of assessment: Metaphors, myths and realities. *Pedagogy, Culture and Society* 15, no. 1: 55–69.
- Taras, M. 2007b. Terminal terminology: The language of assessment. In *Marginality and Difference in Education and Beyond*, ed. M. Reiss, R. Hayes, and A. Atkinson, 52–67. Stoke-on-Trent: Trentham Books.
- Taras, M. 2007c. Assessment for learning: Understanding theory to improve practice. *Journal of Further and Higher Education* 31, no. 4: 363–71.
- Tight, M. 2004. Research into higher education: An a-theoretical community of practice? *Higher Education Research and Development* 23, no. 4: 395–411.
- Torrance, H. 1993. Formative assessment: Some theoretical problems and empirical questions. *Cambridge Journal of Education* 23: 333–43.
- Torrance, H., and J. Pryor. 2001. Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal* 27, no. 5: 615–31.
- Wiliam, D. 2000. Integrating summative and formative functions of assessment. Keynote address to the European Association for Educational Assessment, 9–12th November, Prague, Czech Republic. <http://www.kcl.ac.uk//depsta/education/hpages/dwiliam.html>.
- Wiliam, D., and P. Black. 1996. Meanings and consequences: A basis for distinguishing formative and summative functions of assessment? *British Educational Research Journal* 22, no. 5: 537–48.