

This article was downloaded by: [Michigan State University]

On: 06 March 2015, At: 13:47

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Oxford Review of Education

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/core20>

Formative assessment at the crossroads: conformance, deformative and transformative assessment

Harry Torrance^a

^a Manchester Metropolitan University, UK

Published online: 21 May 2012.

To cite this article: Harry Torrance (2012) Formative assessment at the crossroads: conformance, deformative and transformative assessment, Oxford Review of Education, 38:3, 323-342, DOI: [10.1080/03054985.2012.689693](https://doi.org/10.1080/03054985.2012.689693)

To link to this article: <http://dx.doi.org/10.1080/03054985.2012.689693>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Formative assessment at the crossroads: conformance, deformative and transformative assessment

Harry Torrance*

Manchester Metropolitan University, UK

The theory and practice of formative assessment seems to be at a crossroads, even an impasse. Different theoretical justifications for the development of formative assessment, and different empirical exemplifications, have been apparent for many years. Yet practice, while quite widespread, is often limited in terms of its scope and its utilisation of the full range of possible approaches associated with formative assessment. The paper reviews these issues and explores them in relation to the development of formative assessment in higher education. HE is taken as ‘test case’ of the paradoxical implementation of formative assessment, whereby the aim is, ostensibly, to develop independent and critical learners, while in practice highly conformance assessment procedures are being designed and developed. The paper argues that developers also need to attend to the divergent possibilities inherent in formative assessment, if their aspirations for a more transformative practice are to be realised.

Keywords: formative assessment; divergent assessment; socio-cultural theory; higher education

Introduction

Education as induction into knowledge is successful to the extent that it makes the behavioural outcomes of the students unpredictable. (Stenhouse, 1975, p. 82)

Forget your perfect offering

There is a crack in everything

That’s how the light gets in. (Leonard Cohen, Anthem)

The theory and practice of formative assessment seems to be at a crossroads, even an impasse. Different theoretical justifications for the development of formative

*The Education & Social Research Institute, Manchester Metropolitan University, 799 Wilmslow Road, Didsbury, Manchester M20 2RR, UK. Email: h.torrance@mmu.ac.uk

assessment, and different empirical exemplifications, have been apparent for many years. Yet practice, while quite widespread, is often limited in terms of its scope and its utilisation of the full range of possible approaches associated with formative assessment. Where significant developments have taken place they tend to involve fairly mechanistic forms of activity, to the disappointment of those engaged in facilitating development (cf. Ecclestone, 2002; Marshall & Drummond, 2006; Klenowski, 2009; Swaffield, 2011). Nowhere is this more apparent than in the increasing incursion of a particularly normative version of formative assessment, based on making course objectives and assessment criteria more explicit, into Higher Education. This article will examine the arguments for such developments in Higher Education, taking it as a particular ‘test case’ of the paradoxical implementation of formative assessment. The aim of Higher Education is, ostensibly, to develop independent and critical learners, while in practice highly conformative assessment procedures are being designed and developed. In reviewing this ‘test case’, the paper will also reflect more generally on the tensions and complexities of implementing formative assessment.

Definitions and theories of formative assessment

Definitions of formative assessment essentially revolve around the purpose of assessment being to improve learning. It is argued that the provision of good quality feedback to students during a course on what they have achieved and how they might improve will facilitate learning and improve outcomes. How this might be accomplished remains a matter for considerable debate, however. Formative assessment has been under development in various manifestations for more than 30 years: from mastery learning and criterion-referenced graded assessment programmes in the 1970s and 1980s (Bloom, 1974; Popham, 1978; Pennycuik & Murphy, 1988); through systems approaches to feedback (Ramprasad, 1983; Sadler, 1989); to the variety of approaches to recording, reporting and profiling achievement which emerged in the 1980s and 1990s in the UK (Broadfoot, 1986; Torrance, 1991; Pole, 1993). Current interpretations and exemplifications of formative assessment, also known as ‘assessment for learning’, are often linked to Black and Wiliam’s (1998a) review of the literature, subsequent developments by the UK Assessment Reform Group (2002) and associated work by Black and colleagues (e.g. Black *et al.*, 2003, 2006).

Discussions of formative assessment reflect this genealogy and in many respects reflect the tensions that successive iterations of formative assessment have sought to address. Bennett (2011) reviews various definitions, drawing attention to the recurring focus on purpose—formative assessment being assessment which is intended to provide feedback to both students and teachers on student progress and what more might be done to facilitate such progress. However he also identifies the distinction between whether or not this is conceptualised in terms of the development of diagnostic instruments—i.e. tests used formatively—in the mastery learning tradition, or whether the procedure is seen more in terms of peda-

gogic process and the informal ways in which teachers come to understand student work and seek to assist their learning. Current definitions in the literature favour the latter interpretation. The Assessment Reform Group, for example, (2002) state that:

Assessment for Learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.

However, some argue that such a definition is not actually focused enough on everyday classroom processes and it can too easily accommodate a focus on meeting narrowly defined curriculum goals through coaching for tests (Swaffield, 2011). Addressing this issue, Klenowski (2009) reports the outcomes of the Third International Conference on Assessment for Learning at which it was argued that a 'second-generation definition of Assessment for Learning' (p. 264) was needed (though given the brief history of formative assessment outlined above, in my view this would actually constitute a fifth or sixth generation definition):

Assessment for Learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning. (Klenowski, 2009, p. 264)

Even such more pedagogic definitions still tend to assume that formative assessment must be, and necessarily by definition always is, a positive process, wholly focused on intended learning processes and outcomes, with practice always being congruent with intended outcomes. The argument is that assessment is only formative if it has a positive impact on learning: 'sources of evidence are formative if, and only if, students and teachers use the information they provide to enhance learning' (Klenowski, 2009, p. 264). My own view, on the contrary, is that *all* assessment is formative, of student dispositions and self-identities as learners, as well as of knowledge and understanding, but not necessarily in a positive way. I will return to this issue below, when discussing the notion of *de*formative assessment, but for the moment my point is that current definitions of formative assessment privilege informal classroom processes over more test-based or mechanistic forms of monitoring student progress and providing feedback.

In the United States, early interest tended to revolve around mastery learning and utilising the so-called 'backwash' effect of assessment on the curriculum and teaching methods, to develop programmes of 'measurement-driven instruction' (Popham, 1987; Airasian, 1988). The theory was that if assessment system developers put desired objectives into testing programmes, teachers will teach those desired objectives. As Resnick and Resnick (1992) put it:

You get what you assess; you don't get what you don't assess; you should build assessment towards what you want ... to teach ... (p. 59)

More recently, interest in developing classroom assessment in the USA has merged with debates about formative assessment more generally such that US, and indeed other international interpretations of formative assessment, are now largely co-terminous with UK interpretations (Sadler, 1998; Shepard, 2000; Shavelson *et al.*, 2005; Stiggins *et al.*, 2006; Dywer, 2008; Klenowski, 2009).

Theoretically, the various manifestations of formative assessment are diverse and in key respects stand in contradistinction to each other. Thus the behaviourist tradition of mastery learning takes a very different view of the role of assessment and feedback as compared with more social constructivist perspectives. The behaviourist tradition (and in some respects the systems theory approach as well) insists that we define our objectives, teach to them quite specifically and ensure that teachers and students know what counts as achieving the objective, that is, they know what behaviour is required of them for successful completion of the task. It tends to imply a very structured, hierarchical, 'building block' approach to curriculum organisation and assessment procedures.

More social constructivist approaches see knowledge and understanding as constructed through interaction, rather than transmitted through instruction, placing emphasis on the interaction of teacher and student, student and task, and indeed student and student. The argument, deriving from Vygotsky (1978, 1986) is that it is important to identify not just what the learner has (or has not) achieved, but what they might achieve, what they are ready to achieve, with the help of an experienced teacher or, in some circumstances, a collaborating peer. Thus learning should be 'scaffolded' (Bruner, 1985) and the interactive *process* of assessment is seen to have as much importance for promoting learning as the provision of feedback to students about the product of their endeavours (cf. Gipps 1994, 1999; Torrance & Pryor, 1998; Pellegrino *et al.*, 1999, for a more detailed review of these arguments).

Such social constructivist perspectives are in turn many and varied, a point to which I will return below; but for the moment the issue on which I want to focus is that despite these different theoretical perspectives on formative assessment, which are often manifest in different sorts of empirical instantiations, a basic belief in the efficacy of providing students with formative feedback has become prevalent. While it is the case that different subject communities and different sectors of education systems interpret formative assessment in slightly different ways, reflecting the subject-specific and sector-specific cultures of their communities (Black & Wiliam, 2005; Hodgen & Marshall, 2005), it is also the case that the various theoretical tributaries which have fed contemporary formative assessment debates and practices have largely merged into a rather undifferentiated commitment to the development of 'assessment for learning'. Thus, for example, in parallel with Klenowski (2009), Black and Wiliam (2009) now argue that:

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better

founded, than the decisions they would have taken in the absence of the evidence that was elicited. (p. 9)

So, engaging in formative assessment, and providing formative feedback, is regarded internationally as a rather nebulous ‘good thing’. Yet the empirical evidence of how it is interpreted and implemented in action has disappointed some, while the problems of large scale implementation in an era of test-driven accountability are very significant indeed.

Evidence from practice

Various research studies and reviews have demonstrated the impact of assessment on teaching and learning, along with the efficacy of using feedback to promote learning (see, for example, Crooks, 1988; Gipps, 1999; Harlen & Crick, 2003; Hattie & Timperley, 2007). However, what such reviews have established is that formative assessment *can* improve learning and achievement, not that it will. There is also the issue of how we might judge whether or not the quality of the learning process has improved, and whether or not achievement has been increased. Most studies look at improved test results for evidence and it is probably Black and Wiliam’s (1998a) review which most clearly distilled and disseminated the ‘raising standards’ case for developing formative approaches to assessment. Their associated pamphlet ‘Inside the Black Box’ (Black & Wiliam, 1998b) claimed ‘typical effect sizes of between 0.4 and 0.7 ... larger than most of those found for educational interventions’ (1998b, p. 4). As Bennett (2011) has recently noted, however, the effect sizes quoted in the pamphlet are not directly referenced to previous studies or to any meta-analysis contained in the full review. Bennett observes that:

these effect sizes are not the ‘quantitative result’ ... of the 1998 Assessment in Education review but, rather, a mischaracterisation that has essentially become the educational equivalent of urban legend. (2011, p. 12)

In reviewing these and other studies, Bennett (2011) concludes:

The magnitude of commonly made quantitative claims for effectiveness [of formative assessment] is suspect, deriving from untraceable, flawed, dated or unpublished sources. (p. 5)

Recent individual studies of the effectiveness of formative assessment also indicate, at best, equivocal support with respect to the claims that using formative assessment raises standards of achievement. Thus Wiliam *et al.* (2004) reporting empirical development work deriving from the original Black and Wiliam (1998) literature review, claim that:

despite the cautions noted ... we believe that the results presented here provide firm evidence that improving formative assessment does produce tangible benefits in terms of externally mandated assessments. (p. 63)

Smith and Gorard (2005) argue the reverse. Reporting a small study of the introduction of formative assessment in a school not involved in the Black and Wiliam development programme, i.e. without all the in-service support and possibilities of a ‘Hawthorn Effect’, they conclude that ‘the approach adopted for the treatment group was ineffective overall, and somewhat unpopular with the students as well’ (Smith & Gorard, 2005, p. 37). In a rejoinder Black *et al.* (2005) acknowledge problems with the dissemination of innovation and state that:

We do not claim that formative assessment leads to improved student achievement in *all* classes, with *all* teachers on *all* occasions ... Our claim is that formative assessment in *general* is an effective intervention, although we do not underestimate the difficulties in translating theory into practice. (p. 7, original emphasis)

Stobart (2010) also notes that the teachers and students in Smith and Gorard’s study ‘had only a limited understanding of ... the meaning of comment-only marking’ (p. 151). But this, of course, is precisely the issue—‘translating theory into practice’—scaling up from positive individual studies and general indications of possibilities, or even probabilities, to wide implementation and generally improved practice and achievement.

A further problem concerns whether or not evaluations of the effectiveness of formative assessment should be tied to external measurements of achievement (essentially tests and public examinations) and whether or not increases in such scores, if they are achieved, indicate genuine improvements in educational standards. Much evidence suggests that improvements in test scores do not necessarily indicate real improvements in achievement, because of the problem of coaching and practice or ‘teaching to the test’ (cf. Wyse & Torrance, 2009 for a review of such evidence). Moreover, it is now widely observed that formative assessment is often developed and used precisely in such contexts, i.e. to improve test scores and examination grades, rather than to improve the experience of learning and the quality and diversity of learning outcomes. Thus, for example, Marshall and Drummond (2006) report teachers implementing procedures that reflect the ‘letter’ of formative assessment, rather than the ‘spirit’ which would look to developing students’ understanding and autonomy in their learning experiences. Such findings are reminiscent of similar observations by Torrance and Pryor (1998, 2001) that formative assessment is often ‘convergent’, with teachers focusing on identifying and reporting whether or not students achieve extant curriculum-derived objectives, rather than ‘divergent’ which would be much more oriented towards identifying what students can do in an open-ended and exploratory fashion.

Research in post-compulsory education settings in the UK, including both academic and vocational tracks, has identified widespread formative use of assessment criteria in the context of coaching and practice for external examinations and qualifications (Torrance *et al.*, 2005). Torrance (2007, p. 282) in an article summarising the research reports that:

The clearer the task of how to achieve a grade or award becomes, and the more detailed the assistance given by tutors, supervisors and assessors, the more likely candidates are to succeed. But transparency of objectives coupled with extensive use of coaching and practice ... is in danger of removing the challenge of learning and reducing the quality and validity of outcomes achieved. This might be characterized as a move from assessment of learning, through the currently popular idea of assessment for learning, to assessment as learning, where assessment procedures and practices come completely to dominate the learning experience, and 'criteria compliance' comes to replace 'learning'.

Similar findings have been reported from studies of implementing formative assessment in secondary schools in New Zealand (Hume & Coll, 2009).

So we seem to have reached a moment when developing and implementing formative assessment is generally regarded as a 'good thing', even an orthodoxy of mainstream educational thinking; but much empirical evidence indicates it is difficult to implement consistently and effectively. Moreover, such evidence as does exist on effective implementation suggests that it is the 'specifying objectives and assessment criteria' interpretation, with its resultant emphasis on coaching and practice, which is most easily implemented and observed in practice, rather than what we might term the more ambitious and challenging interpretation of promoting student understanding and autonomy in learning. In other words, despite several decades of debate about the appropriate theoretical underpinnings of formative assessment, and how these might in turn influence its development, it is still a basically behaviourist, objectives-based model which emerges in practice.

Formative assessment in Higher Education

It is with these doubts and cautions in mind that I move to the main focus of this paper—the dissemination and development of formative approaches to assessment in Higher Education. On the face of it, the more ambitious aspirations of formative assessment might be more easily accomplished in the context of higher education—where the aim is to produce graduates capable of independent and critical thinking; where the learners are adults, voluntarily engaging in educational encounters; and where resources for independent and peer-group learning are generally more available than in the compulsory school system. Ironically, however, it is in this most promising of contexts that the basic objectives model still seems to be flourishing and indeed privileged.

A recent special issue of *Assessment and Evaluation in Higher Education* focusing on formative feedback in higher education exemplifies the aspirations for formative assessment in the sector, but also illustrates the limited interpretations of formative assessment which still prevail ('Approaches to Assessment that Enhance Learning': *AEHE*, 35(5), 2010). Most of the articles seem to take that view that assessment in higher education is in need of improvement; that formative assessment should constitute a key element of this improvement; but also that the feedback provided in the context of formative assessment needs more attention from the learner. More interaction is needed between teacher and learner so that feedback is not

just provided, but can be and is acted upon by the student. Such interaction, however, only seems to be understood as working in one direction—students acting on the advice they are given in order to meet the objectives of a course. The editors' introduction states:

This special issue addresses the need to diversify mainstream forms of assessment currently used in higher education ... assessment succeeds when the learner monitors, identifies and then is able to 'bridge' the gap between current learning achievements and agreed goals. It is, however, more questionable whether adequate opportunities are given to the students to be active participants in closing what has been termed 'the loop'. (Hatzipanagos & Rochon, 2010, p. 491)

Closing 'the loop' in this way, however, does rather suggest that students should only pursue the intended learning outcomes of the programme, rather than any others which may arise in the course of the teaching and learning experience.

Additionally, the special issue contains a further introductory commentary from Black and McCormick (2010), drawing attention to the extensive experience of developing formative assessment in the compulsory school system and suggesting the field of higher education could benefit from drawing more on this experience. Yet as we have seen, evidence of effective implementation in the school system is equivocal and, where it does exist, tends towards the mechanistic.

Overall the articles in the *AEHE* Special Issue beg a number of questions about the theory and practice of formative assessment, about how ideas and experience are disseminated within and across substantive fields of intellectual and material endeavour, and about how they are interpreted in different contexts of action. With respect to the core aspirations of higher education, the issue can be stated very bluntly: Are we trying to get students to jump through pre-specified hoops, by making the nature of those hoops more apparent and encouraging students to better understand how the objectives of a course can be met; or are we trying to get students to think for themselves? Additionally, however, the Special Issue begs more general questions about the current 'state of play' with respect to the health of formative assessment, and whether or not we should always strive to produce an 'ideal' solution to a particular problem, i.e. a complete integration of teaching, learning and assessment, rather than live with diverse solutions which may be more or less effective in different circumstances. Is the best way forward to try to plan and construct the perfect functioning system, or should we assume there are inevitably going to be degrees of tolerance which we not only have to acknowledge, but work with as an inevitable and indeed positively beneficial feature of any human system?

Having noted these issues as a backcloth to the remainder of the article, I will now review what I consider to be the main drivers in the development of formative assessment in higher education, highlight some issues and challenges in current theory and practice, and suggest some ways in which the more ambitious aspirations of formative assessment might be pursued. I will not review the *AEHE*

Special Issue in detail, but rather use it as a provocation for the rest of the article, and will return to it as appropriate to the developing argument.

The merger of transparency with formative assessment: conformance assessment

There has been considerable recent policy emphasis, shared by many in HE development roles, on moving the focus of development in the HE sector from teaching to student learning; and from input to output—i.e. intended learning outcomes, including specifying intended learning outcomes and criteria for assessment. This in turn mirrors a more general move towards privileging clarity of outcomes in education systems. The need for assessment to produce more useful information about student achievement has begun to attract attention—more useful information for teachers, for the students themselves, and also for other stakeholders such as employers and government. Norm-referenced, rank ordered degree classifications of 1st, 2.1 or 2.2 do not readily communicate what students have achieved and we have seen a move toward more explicit criterion-based assessment, though degree classifications still persist. Thus a concern for what we might term ‘content standards’, and the production of more useful information about what students can expect to be provided by a course, and come to know, understand and do by the end of their programme of study, has merged with debates about how best to teach, assess and report such content standards. So there has been a move towards greater specification of course programmes, modules and intended learning outcomes, i.e. input; coupled with a new focus on the student experience and what students actually learn, i.e. output:

all courses should be ... designed ... [so] that there should be clear and explicit linkage (alignment) between teaching method and assessment ... so that ... there should be effective and appropriate measurement of the achievement by students of the intended learning outcomes. (Rust, 2007, p. 230, summarising QAA guidance)

In turn, these developments have also been linked to a central role for formative assessment in improving learning and hence the quality of student experience and graduate output:

Formative assessment and feedback should be used to empower students as self-regulated learners [...] self-regulation requires that the student has in mind some goals to be achieved against which performance can be compared and assessed. In academic settings specific targets, criteria, standards and other external reference points ... help define goals [...] Students can only achieve learning goals if they understand these goals, assume some ownership of them, and can assess progress ...

(Nicol & Macfarlane-Dick, 2006, pp. 199–206, my emphasis)

Suggesting that students can ‘*only* achieve learning goals’ in specified circumstances seems to indicate that we know much more about the process of learning than we do. Many people learned many things long before the language of ‘learning goals’ was invented. References to self-regulation can also be understood in the context of the debate about life-long learning and the need for graduates to be flexible learners

over a career or probably (if the rhetoric is to be believed) several careers. Moreover, there are interesting contradictions that seem to be inherent in these arguments. On the one hand higher education seeks to produce flexible self-regulating learners, on the other it is assumed they cannot become such without the learning objectives of a programme being made explicit. How flexible is that?

Furthermore, advice on maximising achievement, or at least measurements of achievement, is often expressed in quite straightforward, instrumental terms, in guide books for undergraduate study:

Make sure that you keep answering the question as written ... The greatest cause of lost marks ... is candidates wandering off on tangents to the question. All the available marks are for points that are directly relevant to the questions. There aren't any more marks for additional knowledge, wisdom or expertise—however brilliant ... (Race, 2007, p. 260)

So there is an agenda of explicitness and transparency couched in terms of fairness to candidates (what they can expect from a course); improving the learning experience (making objectives and assessment criteria very clear); and improving the utility of degree-level education and of graduates.

But, to reiterate this first key point, many of us would argue—including, quite probably, those authors just quoted writing in other contexts—that what Higher Education should, and in many respects does, produce is independent, autonomous learners, capable of understanding and debating the key issues in a discipline. It might also be argued that such independence of thought would be a solid preparation for career progression and life-long learning. However, such nuanced understandings of the qualities which are intended to be developed by tasks and assignments in higher education, and of the quality of particular pieces of work, are not easily captured in lists of goals, standards and criteria. In another contribution to the *AEHE* special issue, Sadler (2010) observes:

Higher education institutions have a responsibility to induct students not only into the mechanics of appraisal but also into a deep appreciation of how complex qualitative judgements can be made ... Qualitative judgements cannot be reduced to a formal set of procedures ... The key... lies in educating students in the art of making ... appraisals in ways similar to those characteristically used by expert assessors ... (Sadler, 2010, pp. 546–547, 536)

The issue then, is that we have moved towards transparency of objectives and assessment criteria, coupled with clear feedback being provided in relation to these criteria, but such a combination of transparency and feedback may not really be considered sufficient to the purpose of higher education. What we have here is not so much formative assessment, but *conformative* assessment.

Is formative assessment always positive anyway? Deformative assessment

As we have seen, many definitions and discussions of formative assessment not only assume that formative feedback will always have a positive impact on the

learner (or neutral impact if the learner doesn't act on the feedback), but actually insist that if such feedback is not used then the activity cannot be categorised as formative assessment (Klenowski, 2009). Sadler (1989), in his now classic exposition of formative assessment, summarises the argument as follows:

... information about the gap between actual and reference levels is considered as feedback only when it is used to alter the gap ... the learner has to (a) possess a concept of the standard ... (b) compare the actual (or current) level of performance with the standard, and (c) engage in appropriate action which leads to some closure of the gap ... (Sadler, 1989, p. 121, original emphasis)

Similarly, to reiterate the definition of formative assessment offered earlier by Black and Wiliam (2009), they argue that:

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited. (p. 9, my emphasis)

There are two problems here, first of all we have the notion of a 'gap' and all that this implies about an incremental, building block view of knowledge. Elsewhere Sadler espouses a much more social constructivist position on the nature of knowledge and learning (e.g. Sadler, 2010), a point to which I will return below. Nevertheless his original paper (Sadler, 1989) has been extremely influential and talk of a gap does imply a very linear and procedural approach to feedback, especially in dissemination and professional development contexts, such as part-time teacher training courses in higher education, where such injunctions may be shorn of their theoretical nuances. We have already noted the misunderstandings that can arise in the school system when new ideas and practices are disseminated without continuing in-service support (viz. the Smith and Gorard (2005) experiment). The idea of 'closing the gap' is all pervasive in the *AEHE* special issue that I have taken as the stimulus for this paper—a special issue that supposedly represents the best of current scholarship on formative assessment in higher education. So questioning the metaphor of the gap is important.

Moreover, not only does the idea of a 'gap' imply a linear model of closure, but it also implies that closure is a good thing—that closure of the gap is what feedback should be trying to achieve. But should it? It is at least arguable, extrapolating from Vygotskian notions of the 'Zone of Proximal Development', noted earlier, that it is in the gap between teacher and learner, and between the learners' present knowledge and future understanding, that the challenge of learning resides—the crack where the light gets in. Furthermore, these are different sorts of gaps, which the metaphor of a single 'gap' does not recognise. The metaphor of a single gap implies that curriculum knowledge is all—the only problem being how to transfer it from teacher to student, how to 'close the feedback loop'. But the issue is not so much to close this 'gap' in any straightforward sense, but to explore and exploit the gaps between teacher and student, and between students' present

and developing understanding through pedagogic action, so that learners come to understand what are the issues at stake, and what learning means for them.

A second problem, particularly exemplified in the extract from Black and Wiliam (2009, p. 9), is the very conscious, rational, deliberative view of decision-making that it implies—as if providing and receiving feedback on students' activities and achievements was some sort of wholly disembodied cognitive process. My argument would be the reverse. Providing and receiving feedback is a highly demanding emotional process, impacting on learners' identities and notions of self-worth. This is particularly the case if the feedback is just a grade or a mark, and especially a low grade or low mark, but it also applies to more extensive forms of feedback. Reading critical comments, even when accompanied by advice about how to improve, can be as disconcerting as receiving a poor grade, especially if the critical comments are unexpected or if the reader disagrees with them. Thus assessment is *always* formative, it will always impact on students and have a central place in what and how students learn, but not necessarily in a positive sense.

It is probably the case that most learners fail most of the time, and many students will come to think of themselves as failures if they do not attain the grades they expect or need, whether they are 'failures' in any real sense or not. This may not be the formative impact we might aspire to, but it happens none the less. Perhaps we should call this *de-formative* assessment. A key issue must be how we minimise such impact if we can. A key starting point must be to acknowledge the nature of the problem.

To be fair to Black and Wiliam, they do recognise issues of affect and attribution. They cite researchers such as Carol Dweck (2000) on the ways in which learners can come to focus on performance goals (i.e. being seen to be doing well on the immediate task at hand, especially in relation to their peers) rather than longer term learning goals. Black and Wiliam (2009) also summarise key issues for the development of formative assessment such as trying to ensure that assessment tasks elicit task involvement rather than ego involvement, and feedback focuses on the task rather than the person. This is easier said than done, however, and a key problem is that their point of reference is the psychological literature on attribution and motivation (Dweck, 2000) rather than ethnographic investigations of the social construction of success and failure in the context of assessment incidents (e.g. Mehan, 1979; Torrance & Pryor, 1998, 2008). Classroom assessment is a public event, and the impact of assessment on individuals has to be understood and addressed holistically and interactively, taking this context into account. Furthermore, attention must be given to the prior social knowledge and expectations that learners bring with them and which will mediate their experience of assessment, as well as the social construction of particular assessment incidents themselves.

Thus assessment will always be formative, but not necessarily in a positive way. A second key issue then is recognising the nature of deformative assessment and how we might better understand and attempt to address the affective impact of assessment.

Socio-cultural theories of learning

One way in which such matters have been acknowledged, at least obliquely, is in recent interest in more complex socio-cultural theories of learning. Without rehearsing all the theoretical tributaries that have contributed to thinking on formative assessment, outlined in the first part of the paper, my argument is that its origins derive from fairly straightforward, behaviourist, stimulus-response theories of learning, particularly in early formulations of mastery learning and the use of various graded objectives schemes of work. This has given way to a more social constructivist position, reflecting Vygotsky's notion of the Zone of Proximal Development:

the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers. (Vygotsky, 1978, p. 86)

It might even be suggested that this now constitutes the main theoretical underpinning of formative approaches to assessment, the reason why feedback is provided, even though it is routinely expressed and realised in terms of intended learning objectives and observable outcomes, rather than better understanding.

A significant issue, however, and one which several papers in the Special Issue of *AEHE* focus on, is student interpretation of and response to such feedback. How is it that some students seem to act on feedback and others do not, that some students seem to 'get it' but others do not? Many of the papers in the Special Issue address this problem from a practical point of view, describing additional ways of trying to communicate and explore assessment criteria with students, including workshops discussing mark schemes, peer assessment and so forth. But we also need to think about the psychological and social processes that we are attempting to engage with when we develop such innovative approaches to assessment, and here, socio-cultural theories of learning, which see learning as a social accomplishment, may be helpful. Socio-cultural theory sees learning more in terms of apprenticeship rather than direct instruction. Thus Lave and Wenger (1991), for example, argue that learning involves induction into a 'community of practice' and the development of learning identities and trajectories over time:

Learning viewed as situated activity has as its central defining characteristic a process we call legitimate peripheral participation ... learners inevitably participate in communities of practitioners and ... mastery of knowledge and skill requires newcomers to move toward full participation in the sociocultural practices of a community ... the meaning of learning is configured through the process of becoming a full participant in a socio-cultural practice ... learning implies becoming a different person [and] involves the construction of identity. (Lave & Wenger, 1991, pp. 29, 53)

Pryor and Crossouard (2008) have developed such ideas into a socio-cultural model of formative assessment. In this, they argue that what one might term the basic core of all approaches to formative assessment—i.e. negotiating understand-

ings of task and quality criteria—should be successively addressed and re-addressed, not just as a decontextualised cognitive activity, but as a socially situated and realised set of interactions, calling attention to the ‘rules of the assessment game’ as well as to the specifics of the topic under discussion.

This theorisation accords with other more descriptive accounts of the complexity of providing usable formative feedback. Thus for example Royce Sadler calls attention to the tacit knowledge that teachers bring to assessment situations; knowledge not just of the specific criteria to apply in a specific situation, but of why they apply, and how they might be set aside in certain circumstances if other criteria come to seem more relevant:

... it is often difficult for teachers to describe exactly what they are looking (or hoping) for ... Teachers’ conceptions of quality are typically held, largely in unarticulated form ... as tacit knowledge. By definition teachers carry with them a history of previous qualitative judgements, and ... the ability to make sound qualitative judgements constitutes a form of guild knowledge. (Sadler, 1989, p. 126)

Determining the quality of complex works requires skilled, qualitative judgements ... judgement is not reducible to a set of measures or formal procedures that a non-expert could apply to arrive at the ‘correct’ appraisal ... students ... need to be inducted into evaluating quality, with ... multi-criterion judgements ... and ... holistic evaluative insights ... (Sadler, 2009, pp.160, 178)

Here we have reference to tacit knowledge, guild knowledge and induction into the history and development of judgement in a discipline. Such concepts parallel very directly the idea of induction into a community of practice and induction into knowledge (Stenhouse, 1975).

However, this process of induction remains problematic. The notion of guild knowledge or tacit knowledge implies the continuation of an opaqueness of judgement which much discussion of criteria insists we dispel. Clearly this is a difficult position to argue in the face of perfectly reasonable demands from students for more clarity of process and procedure. It also presents us with a problem with respect to programme and course design—how much clarification is necessary and how much is enough? A retreat into experienced connoisseurship is not enough; but a constant disinterment and articulation of criteria is too much, and returns us to my opening remarks about the problems of transparency. The discursive nature of the problem is not resolvable in terms of written criteria. Language cannot settle matters of judgment. It can only ever open up more questions (of ambiguity and specificity: ‘But what do you mean by ... ?’). I am reminded of an observation made by Margaret Brown when the National Curriculum was first introduced for schools in England, that criteria have a tendency to ‘multiply like vermin’ (Brown, 1988, p. 19).

The other issue that is raised by the notion of induction into a community of practice is, what sort of community is it that we think we are inducting students into? The theory makes most sense when the activities of the community are agreed and apparent, and the parameters of the community are well understood.

In higher education this would largely refer to disciplinary communities—we are inducting students into the understandings and practices of psychology or sociology or physics, or whatever, with different expectations for the pathways, processes and outcomes of induction (Hounsell & Anderson, 2008). But is this the case any more? Certainly in England the mass expansion of higher education and the changing mission to be much more employment and vocationally-oriented suggests that we are now trying to induct students into a sort of generic gradueness, of data-gathering and problem-solving and evaluating and report-writing, which hardly defines a learning, knowledge-creating community in the sense that Lave and Wenger seemed originally to imply. Moreover the skills of data-gathering and analysis, report-writing and so forth, do not, in and of themselves constitute tools for thinking. Such tools remain embedded in the discourses of knowledge domains and it is only through constant engagement with particular knowledge domains that such tools become apparent. The issue here is epistemological and ontological—what sort of learning identity do we think we are modelling and inducting students into?

Current practice and future prospects: towards transformative assessment

So where are we now, and what are the key elements of current theory and practice that might promise most for development? The following might currently be said to summarise ‘best formative assessment theory and practice’.

We need to try to communicate both the current assessment criteria in play, for evaluating the task at hand, and the process by which these criteria have come to be regarded as important and relevant to a specific task. We also need to convey something of the contingent nature of criteria to the task at hand and the fact that sometimes the most obvious criteria can be put to one side if some new and convincing insight is produced and appealed to. All of this might still be conceptualised as relating to the immediate task of judgement of a particular assignment or project in a particular context.

Additionally we need to try to communicate something of the more generalisable nature of the task at hand, so that feedback carries over not just to the immediate improvement of a project or draft assignment, but to a general understanding of how future assignments might be approached. The notion of meta-cognition comes in here—thinking about thinking—and being aware of how we are thinking, how we are approaching a category of activities and indeed how we come to recognise one activity as similar to or dissimilar from another activity.

So the essential features of a model of current good practice might be said to comprise:

- i) clarifying and thinking about the immediate task at hand: clarifying both ‘task criteria’—i.e. what needs to be done to accomplish the task, and ‘quality criteria’—i.e. what constitutes doing the task well;

- ii) understanding the contingent nature of criteria and the fact that some can outweigh others in particular circumstances;
- iii) meta-cognition—thinking about thinking and the transfer of over-arching criteria to other tasks and situations; for example, the coherence of an argument or the linking of conclusions to evidence.

Points i and ii might be said to constitute the ‘what’ of criteria—what criteria to invoke and employ in a particular setting; iii might be thought of more as the ‘how’ of criteria—how they will always be invoked and embedded in the process of judgement and therefore we should always be aware of the need to explore which criteria are applicable in this new setting or task.

So far, however, even if we accomplish these aspects of formative assessment, we have not really gone much beyond the potentially restrictive parameters of conformance assessment, outlined previously. Attention must also be paid to the divergent possibilities in a learning encounter, to new ways of thinking and new criteria that may be brought into play. Thus, justifying the identification and deployment of different, new criteria must also be countenanced as elements of quality inherent in and above.

Furthermore, a fourth element of the process also needs to be brought into play—that of making both the academic and the social rules of the process of assessment available for scrutiny, including its role in legitimating success, failure and the norms of a meritocratic society. Thus we need an additional element of:

- iv) meta-cognition with respect to the nature of assessment and its legitimating role in the social order.

Point iv might be thought of as the ‘why’ of criteria—why they are invoked to legitimate judgment and therefore why the individual learner should come to understand the process of judgement as social, not personal, with particular criteria, therefore, being open to critique and substitution—if you can come up with better ones, so be it. If you wish to pursue a different learning agenda and can justify this divergence, so be it. The overarching issue to be understood is that the cultural practice of judgement will always be present, but the grounds on which it is exercised can be changed.

At this point, then, we are not just attempting to communicate criteria to learners so that they can understand and deploy them, as given, but also so that they can critique, de-construct and reject them, and understand that other criteria might be more appropriately invoked. This takes us beyond conformance assessment, and ‘criteria compliance’ (Torrance, 2007), towards a rather more transformative vision of assessment which exposes the process of assessment itself to scrutiny, along with the role of criteria in legitimising the discourse of assessment and the social construction of judgement.

Having said this, we also have to recognise that these are rather ambitious, even grandiose, aspirations. For most practical purposes the core of the problem

remains how much criteria clarification is necessary and how much is enough? How can we strike that balance between providing a clear account of the rationale, content and assessment criteria of a programme, while also leaving enough gaps in which the challenge of learning resides and the process of learning can be explored?

Ultimately, however, the solution to the problem is probably less located in answering such questions directly, less located in programme design and the search for a perfectly understood and functioning system, and more located in the vitality and authenticity of the relationships between teachers and students. We need to understand our task as one of collaborating with students to bring about learning, to be alert to the generation of unpredictable outcomes and indeed to regard the production of unpredictable and unintended outcomes as an indication of success, not lack of compliance with the programme. We need to make the rules of the game as apparent as possible, but we also need to try to communicate that we would be happy to see the rules of the game change, if someone comes up with better ones.

Acknowledgements

Anthem: Leonard Cohen Stranger Music Inc., 1992.

Notes on contributor

Harry Torrance is Professor of Education and Director of the Education and Social Research Institute (ESRI), Manchester Metropolitan University. He has undertaken many funded research projects investigating various aspects of assessment and has published widely in the field. He is editor of the forthcoming four-volume *Educational Assessment and Evaluation* (Routledge Major Works series) and is an elected member of the Academy of Social Science.

References

- Airasian, P. (1988) Measurement-driven instruction: a closer look, *Educational Measurement: Issues and Practice*, 7(4), 6–11.
- Assessment Reform Group (2002) *Assessment for Learning: 10 principles*. Available online at: <http://www.assessment-reform-group.org/CIE3.PDF> (accessed 24 August 2011).
- Bennett, R. (2011) Formative assessment: a critical review, *Assessment in Education*, 18(1), 5–25.
- Black, P. & Wiliam, D. (1998a) Assessment and classroom learning, *Assessment in Education*, 5(1), 7–74.
- Black, P. & Wiliam, D. (1998b) *Inside the Black Box: raising standards through classroom assessment* (London, King's College).
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2003) *Assessment for Learning: putting it into practice* (Buckingham, Open University Press).
- Black, P., Harrison, C., Hogden, J., Marshall, B. & Wiliam, D. (2005) Dissemination and evaluation: a response to Smith and Gorard, *Research Intelligence*, 93, 6–7.

- Black, P., McCormick, R., James, M. & Pedder, D. (2006) Learning how to learn and assessment for learning, *Research Papers in Education*, 21(2), 119–132.
- Black, P. & William, D. (2005) Lessons from around the world: how policies, politics and cultures constrain and afford assessment practices, *The Curriculum Journal*, 16(2), 249–261.
- Black, P. & William, D. (2009) Developing the theory of formative assessment, *Educational Assessment, Evaluation & Accountability*, 21, 5–31.
- Black, P. & McCormick, R. (2010) Commentary: reflections and new directions, *Assessment and Evaluation in Higher Education*, 35(5), 493–499.
- Bloom, B. (1974) An introduction to Mastery Learning Theory, in: J. Block (Ed.) *Schools, society and Mastery Learning* (New York, Holt, Rinehart & Winston).
- Broadfoot, P. (1986) *Profiles and records of achievement* (Milton Keynes, Open University Press).
- Brown, M. (1988) Issues in formulating and organising attainment targets in relation to their assessment, in: H. Torrance (Ed.) *National assessment and testing: a research response* (Cheshire, British Educational Research Association), 5–25.
- Bruner, J. (1985) Vygotsky: a historical and conceptual perspective, in: J. Wertsch (Ed.) *Culture, communication and cognition: Vygotskian perspective* (Cambridge, Cambridge University Press).
- Crooks, T. (1988) The impact of classroom evaluation on students, *Review of Educational Research*, 5(4), 438–481.
- Dweck, C. (2000) *Self-theories: their role in motivation, personality, and development* (Hove, Psychology Press).
- Dwyer, C. (Ed.) (2008) *The future of assessment: shaping teaching and learning* (New York, Erlbaum).
- Ecclestone, K. (2002) *Learning autonomy in post-16 education* (London, RoutledgeFalmer).
- Gipps, C. (1994) *Beyond testing* (London, Routledge).
- Gipps, C. (1999) Socio-cultural aspects of assessment, *Review of Research in Education*, 24, 355–392.
- Harlen, W. & Crick, R. D. (2003) Testing and motivation for learning, *Assessment in Education*, 10(2), 169–207.
- Hattie, J. & Timperley, H. (2007) The power of feedback, *Review of Educational Research*, 77, 81–112.
- Hatzipanagos, S. & Rochon, R. (2010) Editorial: introduction to the special issue: approaches to assessment that enhance learning, *Assessment and Evaluation in Higher Education*, 35(5), 491–492.
- Hodgen, J. & Marshall, B. (2005) Assessment for learning in English and mathematics: a comparison, *The Curriculum Journal*, 16(2), 153–176.
- Hounsell, D. & Anderson, C. (2008) Ways of thinking and practising in Biology and History: disciplinary aspects of teaching and learning environments, in: C. Kreber (Ed.) *The university and its disciplines* (London, Routledge).
- Hume, A. & Coll, R. (2009) Assessment of learning, for learning, and as learning: New Zealand case studies, *Assessment in Education: Principles, Policy & Practice*, 16(3), 269–290.
- Klenowski, V. (2009) Assessment for Learning revisited: an Asia-Pacific perspective, *Assessment in Education*, 16(3), 263–268.
- Lave, J. & Wenger, E. (1991) *Situated learning: legitimate peripheral participation* (Cambridge, Cambridge University Press).
- Marshall, B. & Drummond, M.J. (2006) How teachers engage with Assessment for Learning: lessons from the classroom, *Research Papers in Education*, 21(2), 133–149.
- Mehan, H. (1979) *Learning lessons: social organisation in the classroom* (Cambridge, MA, Harvard University Press).

- Nicol, D. & Macfarlane-Dick, D. (2006) Formative assessment and self-regulated learning: a model and seven principles of good feedback practice, *Studies in Higher Education*, 31(2), 199–218.
- Pellegrino, J., Baxter, G. & Glaser, R. (1999) Addressing the ‘two disciplines’ problem: linking theories of cognition and learning with assessment and instructional practice, *Review of Research in Education*, 24, 307–353.
- Pennycook, D. & Murphy, R. (1988) *The impact of graded tests* (London, Routledge).
- Pole, C. (1993) *Assessing and recording achievement* (Buckingham, Open University Press).
- Popham, J. (1978) *Criterion-referenced measurement* (Prentice-Hall, NJ, Englewood Cliffs).
- Popham, J. (1987) The merits of measurement-driven instruction, *Phi Delta Kappan*, 68, 679–682.
- Pryor, J. & Crossouard, B. (2008) A socio-cultural theorisation of formative assessment, *Oxford Review of Education*, 34(1), 1–20.
- Race, P. (2007) *How to get a good degree* (Maidenhead, Open University Press).
- Ramaprasad, A. (1983) On the definition of feedback, *Behavioural Science*, 28, 4–13.
- Resnick, L. & Resnick, D. (1992) Assessing the thinking curriculum, in: B. Gifford & M. O’Connor (Eds) *Future assessments: changing views of aptitude achievement and instruction* (Boston, Kluwer).
- Rust, C. (2007) Towards a scholarship of assessment, *Assessment and Evaluation in Higher Education*, 32(2), 229–237.
- Sadler, R. (1989) Formative assessment and the design of instructional systems, *Instructional Science*, 18, 119–144.
- Sadler, R. (1998) Formative assessment: revisiting the territory, *Assessment in Education*, 5(1), 77–84.
- Sadler, R. (2010) Beyond feedback: developing student capability in complex appraisal, *Assessment and Evaluation in Higher Education*, 35(5), 535–550.
- Shavelson, R., Black, P., Wiliam, D. & Coffey, J. (2005) On linking formative and summative functions in the design of large-scale assessment systems. Available online at: http://www.stanford.edu/dept/SUSE/SEAL/Reports_Papers/On%20Aligning%20Formative%20and%20Summative%20Functions_Submit.doc (accessed 4 July 2011).
- Shepard, L. (2000) The role of assessment in a learning culture, *Educational Researcher*, 29(7), 4–14.
- Smith, E. & Gorard, G. (2005) ‘They don’t give us our marks’: the role of formative feedback in student progress, *Assessment in Education*, 12(1), 21–38.
- Stenhouse, L. (1975) *An introduction to curriculum research and development* (London, Heinemann).
- Stiggins, R., Arter, J., Chappuis, J. & Chappuis, S. (2006) *Classroom assessment for learning* (Princeton, Educational Testing Service).
- Stobart, G. (2010) Making a difference: evaluating the impact of innovations in assessment, in: J. Gardner, W. Harlen, L. Hayward, G. Stobart & M. Montgomery (Eds) *Developing teacher assessment* (Maidenhead, Open University Press).
- Swaffield, S. (2011) Getting to the heart of authentic Assessment for Learning, *Assessment in Education*, 18(4), 433–449.
- Torrance, H. (1991) Records of achievement and formative assessment: some complexities of practice, in: R. Stake (Ed.) *Advances in program evaluation: using assessment policy to reform education* (Greenwich, CT, JAI Press), 231–245.
- Torrance, H. (2007) Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning, *Assessment in Education*, 14(3), 281–294.
- Torrance, H. & Pryor, J. (1998) *Investigating formative assessment: teaching, learning and assessment in the classroom* (Buckingham, Open University Press).

- Torrance, H. & 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.
- Torrance, H. & Pryor, J. (2008) The social construction of success and failure in classroom assessment in England, in: H. Kruger, W. Helsper, G. FoljantyJost, R. Kramer & M. Hummrich (Eds) *Family, school, youth culture* (Frankfurt-am-Main, Peter Lang).
- Torrance, H., Colley, H., Ecclestone, K., Garratt, D., James, D., Jarvis, J. & Piper, H. (2005) *The impact of different modes of assessment on achievement and progress in the learning and skills sector* (London, Learning and Skills Research Centre).
- Vygotsky, L. (1978) *Mind in society* (Cambridge, MA, Harvard University Press).
- Vygotsky, L. (1986) *Thought and language* (Cambridge, MA, MIT Press).
- William, D., Lee, C., Harrison, C. & Black, P. (2004) Teachers developing assessment for learning: impact on student achievement, *Assessment in Education*, 11(1), 49–65.
- Wyse, D. & Torrance, H. (2009) The development and consequences of national curriculum assessment for primary education in England, *Educational Research*, 51(2), 213–238.