Concluding Remarks

Chapter 17

Assessment for Learning: A Compelling Conceptualization

John Gardner

Introduction

At a seminar in 1998, hosted by the Nuffield Foundation at their London headquarters, the Assessment Reform Group launched the Black and Wiliam review pamphlet *Inside the Black Box* (Black and Wiliam, 1996b). The review itself, and the pamphlet, immediately attracted critical acclaim and has continued to enjoy significant impact on assessment thinking throughout the UK and further afield to the present day. However, one moment in the event sticks out clearly in my memory. After the main presentation, a senior educational policy maker stood up and declared that he had heard it all before; we had nothing new to offer. Indicating, with a glance at his watch, that he had commitments elsewhere he promptly left the seminar before the discussion proper got underway. My immediate urge was to rush after him and say 'Yes, you are absolutely right! But it seems to us that, powerful as it might be, formative assessment is actually off the schools' and policy-makers' radar! Surely we need to do something quite urgently if we are to reap the benefits we know are there?' I resisted the urge and instead a year later, at the same venue and with the same sponsors, we injected the urgency we all felt was needed. We launched the pamphlet *Assessment for Learning: Beyond the Black Box* (ARG, 1999). This pamphlet deliberately and directly challenged official complacency and inertia.

Thirteen years on, the Assessment Reform Group has recorded an impressive list of dissemination successes and official endorsements of assessment for learning from, for example, the Scottish and Welsh

governments, the curriculum and assessment agencies of England, Scotland, Wales and Northern Ireland, and from overseas jurisdictions as diverse as Hong Kong and the Canadian province of Alberta. However, in contrast to the situation in Scotland, Wales and Northern Ireland the policy agenda in England remains hamstrung by an accountability agenda which is driving assessment policy. Despite many authoritative, research-based criticisms since 2006, the year of the first edition of this book, I am disappointed to record that schools in England are still being evaluated on the basis of the performance of their students on external assessments. The use of the controversial 'league table', which purports to indicate the relative quality of education in English schools continues unchanged, though abandoned in the three other jurisdictions as divisive and not fit for purpose, that is, forcing schools through competition to raise standards. What it does instead is to increase the emphasis on 'teaching to the test' as schools focus on raising their students' performance in external tests and assessments. There is evidence that the richness of the delivered curriculum suffers and that the pedagogic techniques associated with assessment for learning are neglected.

Paradoxically, assessment for learning's central message, prompted by the research review of Paul Black and Dylan Wiliam (1998a) and disseminated vigorously by the Assessment Reform Group, is that overall standards and individual performance may be improved by actually emphasizing formative assessment techniques such as student self-assessment, negotiation of learning goals and feedback to identify next steps. This message is now strongly established in schools across the UK, though the quality and effectiveness of its implementation is quite another question.

Much progress has therefore been made since the first edition was published, but let me return for a moment to the observations made by our disappointed seminar guest above. I readily concede that the principles and processes of assessment for learning are not novel in any real sense; indeed they have a fairly lengthy pedigree in curriculum and assessment developments in the UK. I could reflect on Harry Black's work with teachers in the early 1980s (Black, 1986) or I could cite the work by Wynne Harlen that led to the publication of professional development materials under the title *Match and Mismatch* (Harlen et al., 1977), to illustrate the point. Such sources would be in keeping with the book's primary focus on schools but I will illustrate the breadth of recognition of the principles we espouse with an example from post-compulsory (vocational) education. The quotation that follows could conceivably have appeared at any time in the last 13 years since the publication of *Inside the Black Box* (Black and Wiliam, 1998b) and the subsequent Assessment Reform Group outputs: *Assessment for Learning:*

Beyond the Black Box (ARG, 1999); *Assessment for Learning: 10 Principles* (ARG, 2002a) and *Testing, Motivation and Learning* (ARG, 2002b).

However, the quotation I reproduce below was actually written in 1986 by Pring as part of an analysis of developments in vocational curricula, initially sponsored by the 1979 publication of the Department of Education and Science's Further Education Unit's *A Basis for Choice*. He argued that a number of implications for assessment had begun to emerge in the wake of the various initiatives in post-compulsory qualifications and summarized them as follows:

> First, *what* had to be assessed was different. A curriculum that stresses personal development, social awareness, cooperative learning, problem solving, is seeking to assess different qualities from those assessed in traditional forms of examination. Secondly, *the purpose of assessment was different*. The main purpose of assessment was the diagnosis of learning needs with a view to promoting the process of learning. It is difficult to provide well-informed guidance, and consequent negotiation of further learning experiences, without some assessment of what the students know or can do. Therefore, it was recommended that the assessment should be part of a continuous, formative profile of the experiences and achievements of the student. Furthermore, it was envisaged that this profile would be the basis of regular teacher/student discussion and guidance of educational progress. The radical difference lies not only in the content of what is taught but also in the processes of learning and thus the demands upon assessment. In its Resources Sheet ... the Joint Board [City and Guilds of London Institute and the Business and Technician Education Council] says:
>
> 'If the individual student is to be enabled to make the most of his/her programme, the quality of the assessment system and its link with supportive guidance will be critical. Most of the assessing will be formative; that is, a regular feedback on performance to the students from all those involved … '
>
> Assessment is thus tied to guidance, negotiation, and the assumption of responsibility for one's own learning. (Pring, 1986:13–14, emphases in original)

There are many such examples, over time, of the acceptance that the classroom assessment techniques comprising assessment for learning are broadly 'good things' to do. However, the specific intention of this book has been to ground this 'goodness' in a credible argument that draws its authority and explanatory power from sound empirical and theoretical contexts. The central arguments have emerged in

various ways throughout the chapters, using research evidence and theory to explain and support the points made. We have attempted to address the specific education-related aspects of assessment for learning, and complementary aspects of summative assessment, but clearly there are many more contextual issues that have a bearing on practice, policy and indeed perception. There are many challenges still remaining for assessment researchers and practitioners, but specifically for policymakers, and in our final pamphlet as a group (ARG, 2009) we drew attention to these. I have summarized three below:

1 Putting effective in-class assessment into practice system-wide;
2 Enhancing confidence in tests and examinations;
3 Justifying the costs of assessment.

Putting effective in-class assessment into practice system-wide

Effective use of formative assessment in English schools is considered to be patchy with teachers, who generally appear to have a strong idealistic commitment to the thinking behind formative assessment concepts, but struggle to put them into practice in the face of competing pressures on their time and priorities. This contrasts with the situation in other countries of the UK. These have reduced some of the critical pressures by rejecting whole cohort testing, as the basis of accountability, and have promoted assessment for learning through rather different kinds of development programmes. For example, the extension of a Thinking Skills and Assessment for Learning development programme in Wales is based on close partnership between civil servants, local authorities and schools, as well as local and national networking to encourage adaptation and spread good practice. Something similar has been effected in Northern Ireland where thinking skills and assessment for learning are firmly embedded in a radical new curriculum. In Scotland, the Assessment is for Learning (AifL) programme has been succeeded by the Curriculum for Excellence programme which has re-affirmed the importance of AfL. As a result, there are high levels of commitment and engagement amongst teachers and learners.

Enhancing confidence in tests and examinations

Assessment data, for the most part based on pupil performance in tests and examinations, are now used in an extraordinary variety of ways, underpinning not just judgements of pupils' progress, but

helping to measure the performance of their teachers, schools and of the nation's education system as a whole, among other uses. These uses can have far-reaching consequences for those being judged by the data. An important question that must continue to be asked is: how reliable is this data? Research presented in this book strongly suggests that we should treat test results with caution. A second important question relates to the validity of tests and examinations: do they measure the aspects of education which society feels it is important to measure? A third issue relates to the impact of publishing information on pupils' test and examination scores. The underpinning accountability agenda lost ground in Scotland, Northern Ireland and Wales and is now largely an English phenomenon, at least in the context of public rankings of schools and their pupils' performances on national tests. Side-effects include the often excessive and inequitable focus of many schools on pupils whose results may be key to a school hitting particular achievement targets; the repetition involved in months of focusing on what is tested and on test practice, which also serves to narrow the curriculum; and the consequent undermining of professional autonomy and morale among teachers. Despite the great volume of material cataloguing these educational side-effects, and the lack of confidence in the reliability and validity of the tests and examinations used, this system persists.

Justifying the costs of assessment

Extraordinary sums of money are now devoted to assessment systems in the UK and the key question is whether these resources could be better spent. In 2005, the consultants PricewaterhouseCoopers published a report (PWC, 2005) based on an investigation carried out in late 2003, which estimated the annual total cost of the English examinations system as £610 million. This total consisted of £370 million which was spent by schools, colleges, awarding bodies and the Qualifications and Curriculum Authority on direct examination costs, and a further £240 million estimated as the cost in terms of staff time in running examination activity in schools and colleges. In Northern Ireland there is a statutory requirement for diagnostic testing of all children in their last four years of primary education and figures released for this (see Gardner, 2010) suggest that in such a small jurisdiction this testing cost over £8,30,000 for 2009–10, not counting teacher time and disruption to class teaching. A report, in 2009, for the examinations regulators for England, Wales and Northern Ireland (OfQual, 2009) collated the incomes of 12 leading awarding bodies covering these countries for the three years to 2007.

These figures gave an average yearly expenditure of £659.3 million, an increase of 15% over the previous two years. We believe that considerable proportions of this money could be better spent on assessments that support learning.

There are many more challenges for the education system but I will finish off with the same example from the first edition, of how one small community espoused assessment for learning; an example that continues to inspire me today.

Assessment for learning: a compelling conceptualization

Any book covering the practice, theory and policy relating to a given educational concept might conceivably claim to provide a comprehensive analysis of that concept. We do not make such a claim for this book on assessment for learning because the extent of existing knowledge and understanding of such a complex process and set of techniques is still in its early stages. We might claim, however, to have assembled an authoritative account of what is known today, however inadequate the extent of this knowledge and understanding might be. Drawing as it does on the work of many researchers and practitioners, as well as our own, this is not an unreasonable claim. We will leave this for others to judge. What we can say categorically about assessment for learning, however, is that it is more often than not a fundamental element of any successful learning context.

Throughout all of the text in this book, therefore, the aim has been to offer what we hope is a 'compelling conceptualization' (Fullan, 2004: 43) of assessment as a process that exists primarily to serve learning. A deep appreciation of this concept was brought home to me very clearly in a presentation I attended on assessment for learning in 2004. The presenters were two teachers, Margo Aksalnik and Bev Hill, from a Rankin Inlet school in the Nunuvut Territory, a new province established in northern Canada in 1999. The main illustration in the talk was of the national symbol of the Inuit people, the Inukshuk. An Inukshuk is a person-like construction of medium-sized rocks, which has been used by the Inuit people for millennia as a means of guiding wayfarers in the treeless and landmark-less expanses of northern Canada. Their various uses include giving directions to good fishing waters or simply reassuring the wayfarer that others have passed the same way, and that they are on the right path. A reproduction of the illustrative model used by the two teachers is presented in Figure 17.1.

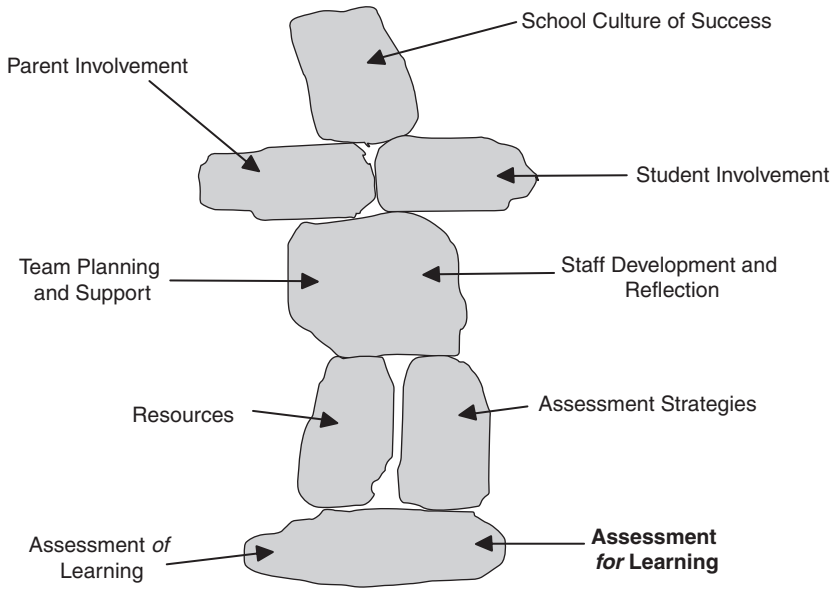

Figure 17.1 An Inukshuk guide to successful education (after Aksalnik and Hill, 2004)

As can be seen, they placed assessment for learning squarely in the set of main ingredients designed to create a school with a culture of success. The other elements included teachers, their planning of the learning activities, their teaching and assessment strategies, their capacity to reflect about their own and their students' learning, and the resources they bring to the learning environment. Outside of the classroom, additional elements include professional development and team support for the teachers while outside of the school, the positive involvement of parents adds to the recipe for success.

It is arguable that other aspects of a successful school could be found to populate the Inukshuk's frame – successful sporting programmes or a students' council, for example. No doubt they and other features of successful schools are also somewhere within the model, but the community-based context in which the two teachers introduced assessment for learning to their school dispelled any notion that its inclusion in the Inukshuk was either whimsical or contrived for the event (a seminar on assessment for learning on Vancouver Island). They recounted that: 'The Elders met to consider these new approaches and had the concept of assessment for learning explained to them. They then came up with a word to identify the dynamic – the resonance – between teaching, learning and assessment. (Aksalnik and Hill, 2004)

This new word, in the Inuktitut language of the Inuits, is:

and is written in Roman form as Illitaunikuliriniq (or in sound form: ee-lee-tau-nee-qu-lee-ree-nee-kay).

Most non-Inuit educationalists will have difficulty articulating this word but they will not fail to empathize with the assessment for learning aspirations of this small community in Canada's frozen north.

References

Aksalnik, M. and Hill, B. (2004) Oral presentation to 'Making Connections', Assessment for Learning Symposium, Vancouver Island, Classroom Connections, Courtenay, July.

ARG (1999) *Assessment for Learning: Beyond the Black Box*. University of Cambridge: Assessment Reform Group.

ARG (2002a) *Assessment for Learning: 10 Principles*. University of Cambridge: Assessment Reform Group.

ARG (2002b) *Testing, Motivation and Learning*. University of Cambridge: Assessment Reform Group.

ARG (2009) *Assessment in Schools: Fit for Purpose*?, available at www.tlrp.org/pub/documents/assessment.pdf

Black, H. (1986) 'Assessment for learning', in D.J. Nuttall (ed.), *Assessing Educational Achievement*. London: Falmer. pp. 7–18.

Black, P. and Wiliam, D. (1998a) 'Assessment and classroom learning', *Assessment in Education*, 5: 7–71.

Black, P. and Wiliam, D. (1998b) *Inside the Black Box: Raising Standards through Classroom Assessment*. London: King's College (see also *Phi Delta Kappan*, 80: 139–48).

Fullan, M., Bertani, A. and Quinn, J. (2004) 'New lessons for districtwide reform', *Educational Leadership*, April: 42–6.

Gardner, J. (2010) Report from the Working Group on the InCAS Errors of October 2009, Department of Education for Northern Ireland, available at: www.deni.gov.uk/incas_working_group_report.pdf

Harlen, W., Darwin, A. and Murphy, M. (1977) *Leader's Guide: Match and Mismatch Raising Questions and Match and Mismatch Finding Answers*. Edinburgh: Oliver and Boyd.

Ofqual, Welsh Assembly Government, Council for Curriculum, Examination and Assessment (2009) *Annual Qualifications Market Report, 2009*. Ofqual.

Pring, R. (1986) 'The developing 14–18 curriculum and changes in assessment', in T. Staden and P. Preece (eds), *Issues in Assessment, Perpectives 23*. Exeter: University of Exeter. pp. 12–21.

PricewaterhouseCoopers (PWC) (2005) *Financial Modelling of the English Exams System* 2003–4.