

## Chapter 1

---

# Introduction

*John Gardner*

On first inspection, the title of this book arguably places learning, one of the most fundamental processes in a person's lifecourse, secondary to one of the most contrived processes, the assessment of that learning. Our intention, however, is quite the opposite. As members of the Assessment Reform Group (ARG) we have accumulated over 21 years of collective research into assessment policy and practice, and many more years as individuals. As with the first edition, this second edition of *Assessment and Learning* is a book that unapologetically places learning at the centre of our concerns while unambiguously underscoring the importance of assessment in that learning.

The Assessment Reform Group was a research and lobbying group based in the UK and though it is natural for us to turn to our own contexts to illustrate analyses of assessment practice, the key aspiration throughout the group's existence has been to collate and use research from around the globe to develop a better understanding of how assessment can significantly contribute to learning. The reader will therefore find a liberal sprinkling of research-informed insights from a wide variety of international contexts. Here and there, throughout the book, we refer to various types of learning contexts in these countries but it is fair to say that, with the notable exception of Kathryn Ecclestone's Chapter 9, we draw heavily on the compulsory phases of education (roughly 4–16 years in most countries) to contextualize the practice of assessment for and, in some cases, of learning. It is also fair to say that it is in this context that the majority of research and experimentation has been recorded. We recognize that it is beyond the capacity of any one book to cover the huge span of educational endeavour in a world in which lifelong learning is the name of the game but we hope that the concepts and processes we illuminate throughout the chapters, such as learner engagement,

feedback, motivation and pedagogic style, are key to any learning environment facilitated by teaching or instruction. Translating them to other learning contexts such as work-based learning, adult and community education or post-compulsory education, is not straightforward but the principles and practices will be relevant.

In most developed countries, the pursuit of reliable and valid means of assessing people's learning generates high volumes of published discourse and, not infrequently, dissent; the documentation on the various assessment policies, practices and theories could conceivably fill whole libraries. Some of the discourse and much of the dissent relate to whether the use to which assessment is put is valid, or, to put it more mundanely, useful to the learners themselves or to other audiences. Our pursuit is somewhat different. We would argue that learning should take centre stage and we address the role that assessment should play in this. Assessment is our focus but learning is the goal – and the implication for this edition as with the first is that in much of what follows it is assessment *for* learning that dominates the writing with assessment *of* learning given its secondary but still important place.

Two phrases, 'formative assessment' and 'assessment for learning', are used throughout all of the chapters that follow. The older phrase, 'formative assessment', can be traced back to Scriven's (1967) concepts of formative and summative evaluation, distinguished at the time solely on the basis of when the evaluation in question is carried out. While timing is merely one of the distinctions today, formative assessment remains a widely used concept in education. However, it is sometimes used to describe a process in which frequent ad hoc assessments, in the classroom or in formal assessment contexts such as practical skills work, are carried out over time and collated specifically to provide a final (summative) assessment of learning. Such assessments potentially do not contribute to the students' learning. The second phrase, 'assessment for learning', came into use in the late 1980s and early 1990s (the chronological sequence and provenance is set out in detail in Siobhan Leahy and Dylan Wiliam's chapter 4) and may therefore be considered a somewhat 'newer' concept.

In truth, though, assessment for learning comprises the same time-honoured practices as formative assessment, that is, 'what good teachers do and have always done' (ARG, 2002) when using assessment to assist students to take the next steps in their learning. In contrast to the term 'formative assessment', however, assessment for learning is arguably less likely to be used to describe the summative use of multiple assessments of learning. The words focus squarely on the essence of our pursuit: the promotion of assessment to support learning and this is neatly contra-distinct from assessment *of* learning. In the

final analysis there is little of substance to distinguish the two terms ‘formative assessment’ and ‘assessment for learning’, but for the wider educational and policy-making audiences we feel that the latter is more accessible than the more technical term, ‘formative assessment’. That said, we are content to use both phrases interchangeably, when there is no ambiguity in the type of assessment process being described.

In order to ensure we remain consistent in how we describe the type of process that assessment for learning is, we have defined it to be: ‘the process of seeking and interpreting evidence for use by learners and their teachers, to identify where the learners are in their learning, where they need to go and how best to get there’ (ARG, 2002).

Unpacking this deceptively simple definition, in terms of classroom practice, reveals a complex weave of activities involving pedagogic style, student–teacher interaction, self-reflection (teacher and student), motivation and a variety of assessment processes. For example, teachers need to plan the learning environment and activities, students need to engage in the assessment of their learning and teachers need to assess the extent of the students’ understanding as they are learning. They then need to challenge and support these students to enable them to reach the next stage in their learning progress. An analysis of such a complex learning approach could never be exhaustive but we have tried to make it accessible through a previous publication entitled *Assessment for Learning: 10 Principles*. These principles are mentioned in various places in the chapters that follow and are summarized below:

#### *Assessment for learning*

- Is part of effective planning;
- Focuses on how students learn;
- Is central to classroom practice;
- Is a key professional skill;
- Is sensitive and constructive;
- Fosters motivation;
- Promotes understanding of goals and criteria;
- Helps learners know how to improve;
- Develops the capacity for self-assessment;
- Recognizes all educational achievement. (ARG, 2002)

All of these qualities, which we attribute collectively to assessment for learning, appear in various guises throughout the book – in their practice, in the theories underlying them and in the educational policies that relate to them.

## Purposes and practice, impact, theory and validity and reliability

Under these generic headings the structure of the book proceeds in four parts, which in turn address the purposes and practice of assessment (Part 1), its impact in a variety of contexts including its impact on learners (Part 2), its theoretical underpinnings (Part 3), and the enduring issues of validity and reliability (Part 4).

Part 1 – Purposes and Practice – represents a significant enhancement of the book with three new chapters. The first chapter (Chapter 2) by Paul Black and Dylan Wiliam draws as before on empirical evidence from several projects to portray how the findings from formative assessment research may be translated to classroom practice. In addition to significant changes to pedagogy, the chapter demonstrates that a full espousal of assessment for learning creates inevitable changes in teachers' and learners' understanding and attitudes to learning itself. Dave Pedder and Mary James pick up the story in Chapter 3 as they did in the first edition. This chapter draws on another, *Learning How to Learn*, to tackle how such changes can be promoted and supported through teachers' professional learning. Central to the findings is the observation that very often, in an assessment for learning context, there is little to distinguish between the processes of learning for students and teachers. Chapter 4 is a new chapter from Siobhan Leahy and Dylan Wiliam that proposes up to 80% increases in students' speed of learning when formative assessment practices are truly integrated 'minute-by-minute and day-by-day' into teachers' classroom activities.

Chapter 5 is also new and draws on ARG's Assessment of Significant Learning Outcomes, the ASLO project. Richard Daugherty and colleagues present case study evidence that takes the analysis of the relationship between curriculum and assessment beyond the simple notion of explicit outcomes of assessment being in some way aligned to, or congruent with, a pre-specified curriculum. Rather than thinking in terms of aligning assessment more closely to curriculum, they argue that the construction of learning outcomes is better understood as a complex, non-linear, interacting system that embraces curriculum, pedagogy and assessment.

In Chapter 6 Wynne Harlen explores whether it is feasible for evidence gathered for one purpose to be used for another, focusing specifically on formative and summative assessments. Proposing a spectrum of possibilities that allows an overlap of purposes in its middle region, the chapter highlights the pitfalls of blithely advocating dual usage of the same information and suggests the conditions under which its integrity may be preserved. The final chapter in the section, Chapter 7 by John Gardner, tackles the thorny issue of what

is considered to be quality in assessment practice. This chapter draws on the Analysis and Review of Innovations in Assessment project (ARIA) to identify key principles for quality assessment and then uses a selection of these to discuss the central aim for all assessment, the challenges of assessing complex learning, assessment and social inequality, and the public understanding of assessment.

Part II – Impact – focuses on the impact of assessments in several areas: on learners themselves, in vocational education and across several international contexts. Chapter 8 is a new chapter by Louise Hayward, in which she reflects on what we might learn from listening to learners. The chapter also identifies the process of working with learners, to explore the world as they experience it, as a research priority for the future development of assessment for learning. In Chapter 8, Kathryn Ecclestone probes trends in UK vocational education assessment that have resulted in what she terms the ‘Holy Grail of crystal clarity’. These trends are skilfully unpacked as she identifies the pressing need to explore educational questions about the sort of person our current learning and assessment culture is creating, and the forms of knowledge, skills and dispositions that assessment fosters, overlooks or discourages. Judy Sebba’s Chapter 10 acknowledges the problems inherent in any attempt to compare practices across cultures but points to the commonalities of understanding and practice across several nations. Wynne Harlen’s Chapter 11 concludes the section by addressing a key element of formative assessment’s armoury: promoting learners’ motivation. Drawing on research from around the world, she argues that some summative assessment practices may have a negative impact on learners, while steps that ameliorate the worst effects, by developing and sustaining learner motivation, are often based on the principles of assessment for learning.

Part 3 – Theory – begins with a revamped chapter from Mary James (Chapter 12), which no longer reviews a history of learning theories (as in the first edition) but instead focuses on the problems and possibilities of developing assessment practice that is congruent with socio-cultural learning theory. This ‘third generation’ learning theory is neatly exemplified from a classroom perspective in a contribution from Jenny Lewis. Paul Black and Dylan Wiliam return in Chapter 13 with a fully revised and updated chapter offering a clear rationale for formative assessment, insights into its theoretical basis and a theoretical framing of formative assessment activities designed to suggest ways they may be evaluated and further improved.

Part 4 – Validity and Reliability – presents three authoritative chapters on these central themes of assessment practice. Gordon Stobart’s Chapter 14 picks up the validity issue and concludes that there is a simple test for the validity of any assessment for learning process: did

learning take place as a consequence? He poses a consequential validity argument and focuses on all aspects of the cultural and learning context, the quality of classroom interaction, the teacher's and learners' clarity about what is being learned and the effectiveness of feedback. Paul Black and Dylan Wiliam's chapter 15 is a significant updating of their first edition chapter, and carries on a well-informed and well-argued debate about the perceived reliability of national and other external (to the school) testing. They conclude with the hard-hitting statement that trust in test scores is clearly not justified and is a serious issue when decisions on a student's future may be based on this misguided trust. The final chapter in this section, Chapter 16, is a new one from Paul Newton, which returns to the overarching keystone in assessment practice: validity. Taking a 'purposes' approach, the chapter warns the unwary about inappropriate (invalid) uses of assessments and points to the need for a sophisticated conception of validity to ensure assessment results are 'recycled' responsibly.

Chapter 17 from John Gardner rounds up the book with a concluding discussion on its main messages.

## A note on the Assessment Reform Group

From its establishment in 1988, as the then Assessment Policy Task Group of the British Educational Research Association (BERA), the Assessment Reform Group occasionally changed in personnel but doggedly pursued the agenda of improving assessment in all of its forms. The founding members were Patricia Broadfoot, Bryan Dockrell, Caroline Gipps, Wynne Harlen and Desmond Nuttall, and its first task was to consider the implications of the 1988 Education Reform Act. Following Bryan Dockrell's retirement, Mary James and Richard Daugherty joined the group in 1992 and, in 1994, after the untimely death of Desmond Nuttall the previous year, Gordon Stobart and John Gardner also joined.

The membership then remained more or less unchanged until Caroline Gipps and Patricia Broadfoot moved on, in 2000 and 2002 respectively. Very able replacements were on hand and when the group finally wound up in 2010, it included Jo-Anne Baird, Paul Black, Kathryn Ecclestone, Louise Hayward and Paul Newton. Judy Sebba and Dylan Wiliam had also been members in the interim period before external commitments took over their time. In this edition we are very pleased to welcome them back along with guest contributors: Dave Pedder, Siobhan Leahy and Jenny Lewis.

From 1997, when BERA ceased to sponsor policy task groups, the ARG worked as an independent group with funding primarily from

the Nuffield Foundation. The Foundation generously supported a variety of our activities including group meetings, regional seminars and the dissemination of our work. This funding was crucial to our success and we would be very remiss if we did not take the opportunity here to acknowledge our grateful appreciation of both BERA's and, in more recent times, the Foundation's support.

On that note, this introduction to the second edition would be seriously deficient if acknowledgement of our sources and influences was not formally recorded. Over the period of its existence, the group worked with many people including teachers, academics and curriculum and assessment agency personnel from around the world, local authority advisers and district superintendents, government officials, politicians and, most importantly, students in a variety of national and local contexts. There are too many to name and it would be inappropriate to single out specific people. However, the content of this book has been influenced by them all and we humbly record our thanks to everyone with whom we have had the privilege to work.

### *Postscript*

In March 2008 members of the ARG, using funds from their publications, spent 24 hours 'away from it all' at Ross Priory on the shores of Loch Lomond. The aim was to take stock of past work particularly on formative assessment, to consider implications of current assessment policy and to decide the future for the Group. The decision was made to wind up the Group in 2010–21 years after it began. More than half of the group is now retired, though all of them continue to be active in assessment research. On 8 June 2010, Cambridge Assessment hosted an event to say 'thank you and farewell' to the ARG on the occasion of our retirement. Held at Downing College, Cambridge, it provided an opportunity to reflect on the contribution that the Group has made to assessment policy and practice, both in the UK and beyond. With a view to the future, members of the Group offered reflections on what makes for successful assessment reform, encouraging a broader debate on the topic. Materials from the event, including a write-up of proceedings by education journalist Warwick Mansell and photos from the occasion, can be found at <http://www.assessnet.org.uk/arg/>.

### **References**

- Scriven, M. (1967) 'The methodology of evaluation', in R. W. Tyler (ed.), *Perspectives of Curriculum Evaluation*. Chicago: Rand Mc Nally. pp. 39–83.

