Assessment for Learning *in* Higher Education

EDITED BY
PETER KNIGHT

S E D A

Assessment
for
Learning
in
Higher
Education

The Staff and Educational Development Series
Series Editor: Sally Brown
Assessing Competence in Higher Education Edited by Anne Edwards and
  Peter Knight
Assessment for Learning in Higher Education Edited by Peter Knight
Research, Teaching and Learning in Higher Education Edited by Brenda Smith
  and Sally Brown

SEDA is the Staff and Educational Development Association. It supports
and encourages developments in teaching and learning in higher
education through a variety of methods: publications, conferences, net-
working, journals, regional meetings and research – and through the
SEDA Fellowship Scheme. Further details may be obtained from:

The SEDA administrator
Gala House
3 Raglan Road
Edgbaston
Birmingham B5 7RA
Tel: 0121-440 5021
Fax: 0121-440 5022

Assessment for Learning

in

Higher Education

EDITED BY
PETER KNIGHT

Contents

Notes on Contributors

Madeleine Atkins is Senior Lecturer in Educational Management at the University of Newcastle upon Tyne and Head of the Department of Education. She has a long-standing interest in research on teaching and learning in higher education and in staff development. With George Brown she wrote *Effective Teaching in Higher Education.*

David Boud is Professor of Adult Education at the University of Technology, Sydney. He is a long-time proponent of the need to appraise academic practices, including assessment, in terms of how students see them and of their impact upon student learning. Amongst his many publications are *Appreciating Adults Learning, Reflection: Turning experience into learning, The Challenge of Problem-based Learning* and a large number of articles, including ones on assessment and self-assessment.

Sally Brown is based in the Educational Development Service at the University of Northumbria at Newcastle. She is Chair of the SEDA publications committee and runs staff development workshops in the UK and overseas. She has written extensively on educational development and teaching issues. Recent books include *Assessing Learners in Higher Education* and *Research, Teaching and Learning in Higher Education.*

Phyllis Creme is involved both in educational development and in Film Studies teaching at the University of North London. She is based in the university's Centre for Higher Education and Access Development and has responsibility for teaching and learning innovation.

Kate Day is the Research and Development Officer and **Dai Hounsell** is the Director of the Centre for Teaching, Learning and Assessment at the University of Edinburgh. They work collaboratively in a number of areas, including feedback and evaluation, quality assurance, and assessment. Dai is coordinator for a large-scale project on assessment strategies in Scottish higher education, in which Kate is involved as a Project Director.

T Dary Erwin is Director of the Office of Student Assessment at James Madison University, Virginia and Professor of Psychology. His published works include the recent *Assessing Student Learning and Development.*

7

Nancy Falchikov works in the Department of Social Sciences at Napier University, Edinburgh. In addition to self- and peer-assessment, her research interests include the effects on student learning of matches and mismatches in the learning environment in terms of student approaches to learning, methods of teaching and assessment practices.

Bob Farmer and **Diana Eastcott** work in the Learning Methods Unit at the University of Central England in Birmingham. They are both practising teachers, trainers, educational consultants and authors of a wide range of articles on teaching and learning.

Hazel Fullerton has used her design training in posts in video production and training; in setting up resource-based learning in schools; as a communications tutor in Further Education; and as a technology tutor/counsellor with the Open University. She is now Academic Staff Development Coordinator at the University of Plymouth.

Romla Hadrill works at the Manchester Metropolitan University in the field of post-compulsory teacher education. Current work includes consultancy for the CVCP Staff Development Unit on a project that is mapping higher-level NVQs and the work of staff in higher education. She is vice-chair of the Universities Professional Development Consortium for post-compulsory education and training and a lay inspector for the Further Education Funding Council.

Irene Harris is Senior Lecturer in the Faculty of Management and Business at the Manchester Metropolitan University. Her interest in assessment issues has been fuelled by her work with students, her own experience as a mature student and by being a manager of an Enterprise in Higher Education project.

Peter Knight is a lecturer in Lancaster University's Department of Educational Research.

Ivan Moore is currently Assistant Director of Educational Development at the University of Ulster. A Fellow of SEDA, he acts as a consultant on teaching, learning and educational development, about which he has also written three books and a number of papers.

Phil Race is Professor of Educational Development at the University of Glamorgan. His interest in assessment stems from many years as an examiner and from work helping students to improve their learning strategies and communication techniques. An active author and consultant, his most recent books include *500 Tips for Students* and *500 Tips for Tutors*.

Chris Rust is Principal Lecturer in the Educational Methods Unit of Oxford Brookes University. Through the Oxford Centre for Staff Development, he regularly runs workshops around the country on issues such as assessment. He is co-author of *Strategies for Diversifying Assessment* and has edited two successful induction packs for new lecturers.

Note: Many of the publications mentioned here were written by more than one author. To keep this snappy, I've generally not mentioned these co-authors.

Preface

The assessment of student learning has often been seen as a tiresome and harmful necessity. Tiresome, because of the amount of work it imposed upon learners and tutors and because it seemed to get in the way of worthwhile learning; harmful because it seemed to encourage cramming, superficiality and conformity; and a necessity because without it there was no way for universities to show that they maintained high standards. Besides, without assessment, what was there to make students work?

An alternative view has emerged in schools and higher education, namely that 'student assessment is at the heart of an integrated approach to student learning' (Harvey, 1993, p.10). It is becoming appreciated that assessment arrangements can be diverse; can support ambitious curriculum aims; and can foster understanding. So far from there being a tension between assessment and learning, reformed assessment arrangements might be a necessary condition for better student learning to take place.

In this belief, the Staff and Educational Development Association (SEDA) organized a conference in May 1994 around the theme of assessment for better student learning. The chapters in this collection constitute a selection of the 32 conference presentations. All have been rewritten for this book, although Dary Erwin's chapter (Chapter 3) was especially written for this book.

While I have dealt with the mechanics of editing, I am indebted to colleagues for their academic help in planning the collection. First comes Sue Drew, of Sheffield Hallam University, who was a vigorous and efficient conference organizer, leading the SEDA conference planning committee. That committee also comprised Joyce Barlow of the University of Brighton, Richard Kemp of the University of Glamorgan and Ranald MacDonald of Sheffield Hallam University. It benefited from the support of Jill Brookes, SEDA administrator and Jessica Claridge of the University of Exeter. At the conference I was grateful to the following for their advice: Dr Liz Beaty of the University of Brighton, Ms Sally Brown of the University of Northumbria at Newcastle, Professor Arnold Goldman of the University of Kent, Professor Phil Race of the University of

Glamorgan and Mr Chris Rust of Oxford Brookes University. Three of these people have papers in this collection but they were not party to their selection.

Peter Knight, March 1995

Introduction

Peter Knight

ASSESSMENT – EVIDENCE OF QUALITY

Assessment is a moral activity. What we choose to assess and how shows quite starkly what we value. In assessing *these* aspects of chemistry or by assessing German in *that* way, we are making it abundantly clear what we value in *this* programme and in higher education in general. So, if we choose not to assess general transferable skills, then it is an unambiguous sign that promoting them is not seen to be an important part of our work and of our programme. That position is, of course, defensible in several ways, not least on the grounds that these skills (whatever 'skills' might be) may not be quite so general, let alone easily transferable (see Eraut, 1994b and Atkins, this volume). Yet, whether the intellectual position is defensible or not, in choosing not to assess learners' general transferable skills through these programmes we reveal our values.

And we reveal them quite starkly. In writing a mission statement, a programme plan or a validation document, skilled drafting allows us to lay claim to a wonderland of concepts, skills, competences and the like, of which our students are to be made citizens. But for those who want to know about the quality of a course, programme or institution, the test is whether these goals are assessed and how well they are assessed. In a sense, the way students are assessed is the 'DNA evidence' of their learning experiences. We might say that we have been trying to promote these skills, understanding in that area, or competence in this element but if there is no evidence of appropriate assessment, then the DNA evidence belies the claim. At best, the absence of assessment suggests that our intentions have not been completely realized. At worst, it says that our intentions were rhetorical, for the benefit of auditors, not students.

To illustrate some of the things that might be inferred about a university and its attitude to student learning from assessment data, I want to look at the University of Arcadia, which could be in Wigan, Weymouth or Wolvercote. Fictional though the university is, the data come from real, unpublished case studies.

ARCADIA AND ASSESSMENT

At Arcadia the same degree may be gained through courses offered by different teams which work on different sites. The courses have very different assessment requirements. It has only been possible to look at written work that contributes to degree classification, but the picture of assessment demands is quite striking, as Table I.1 shows.

At the very least there is a moral issue here about the differential requirements upon learners who will all end up with a University of Arcadia degree. This unease about the fairness of the system might be compounded by data showing that in the year in which these data were collected, 60 per cent of students on BA course 3 graduated with an upper second or first class honours degree, even though they had

Table I.1   *A comparison of assessment requirements for some post-Year 1 students at the University of Arcadia*

Degree title	Number of coursework items	Total length of coursework (words)	Dissertation requirement (words)	Length of examinations (hours)
Education, subject 1, course A	18	43,000	optional	24
Education, subject 1, course B	23	50,000	compulsory, 10,000+	17
Education, subject 2, course A	Not clear in validation documents	Not clear in validation documents	optional	21
Education, subject 2, course B	Not clear in validation documents	50,500	compulsory, 10,000+	12
BA, subject 1, course A	10	21,000	compulsory, 10,000+	13
BA, subject 1, course B	22	40,000	compulsory, 10,000+	16
BA, subject 2, course A	16	32,000	compulsory, 10,000+	16
BA, subject 2, course B	13–15	39,000	optional	18

entered the course with modest 'A' level scores (a mean of just under 11 points). Students entering another BA course had lower mean 'A' level scores (just over nine points) but only 13 per cent of them got an upper second or first class honours degree. Those entering with the best 'A' level scores took yet another BA course but just 42 per cent got an upper second or first class honours degree. This odd pattern cannot be simply attributed to the assessment procedures in use, for the quality of teaching might have played an important part.

A closer examination of validation documents drawn from a sample of eight departments at Arcadia shows further diversity between departments, this time in variations in the balance between assessment through coursework and by examination; in the amount of assessed work required by different departments; in the form of assessment items; and in the timing of assessment. Students taking joint honours degrees would often be assessed quite differently in the two strands of their degree. What the validation documents had in common was a reliance on a narrow range of assessment methods; silence about assessment criteria; an absence of plans to assess 'core competences' or 'general, transferable skills'; and a general indifference to self- and peer-assessment.

So, what defines a degree from this ubiquitous university? The answer seems to be that only indifferent assessment practices cut across the proliferation of diversity. Survival of the fittest has not operated here: rather, perhaps, survival of the flattest – those assessment arrangements which are least likely to jolt the passage of traditional, well-oiled teaching and learning juggernauts.

A study of students at Arcadia showed that departmental codes of practice had a lot to say about the bureaucracy of assessment but little about the purposes and criteria of assessment and that,

> students often don't know why the system is as it is, or how they are meant to do something. Basic questions remain unanswered, for example, 'What skills am I being assessed on?', 'Why do we have exams?' Students have numerous doubts regarding the *reliability, validity and effectiveness* of assessment, as well as the degree to which it *contributes to the learning process.*

Another study found that students doing dissertations were not sure about the purpose of doing so and were often quite intimidated by the process of coping with this form of assessment. A third study found that the quality of feedback on assessed work left much to be desired, especially since it was usually slow to arrive and so bound to the specifics of the task in hand that it failed to offer much useful, general advice for doing better on the next task.

These data would be consistent with a view that at Arcadia assessment was a vital ritual in the maintenance of some hazy features of the social order. They would not seem to be so consistent with the common sense view that assessment was designed to buttress students' learning. If that is so, it would appear that Arcadia is ripe for an anthropological study.

Examination of the marks attained by students on each unit of assessment shows some interesting features. In each of three non-natural science departments studied, coursework marks are higher than examination marks across a two-year period. The largest mean difference is nearly seven marks and the smallest is just over one. Clearly, the assessment system is working against those who are better at examinations than at coursework, which may be desirable but which also is an expression of certain educational values.

Those who chose to take examinations during the second year had mean scores that were about 1.5 per cent lower than their third-year examination marks, although coursework marks for second-year work were not statistically significantly different from third-year coursework marks. This might make us wonder about the wisdom of modularization with its emphasis on completing unit assessment when the unit itself is completed. It might also make us wonder about the claim that students develop in their third year. If their coursework improves, then these marks do not show it!

Similarly for data showing that women outperform men in these subjects: their mean coursework score is nearly four marks higher than the men's, while the mean examination score is nearly three marks higher. Is this to be interpreted as showing that the assessment system is unfair to men, or that teaching and learning do not sufficiently engage men, or that women taking those subjects in those years were simply cleverer?

Examinations have a long history and it might be assumed that a certain objectivity attaches to the information they provide. However, at Arcadia we can see variations in the performance of different groups of students that would seem to need some reasoned justification. None appears to be to hand, although Erwin, in Chapter 3 of this volume, is clear that in a North American setting such variations would be probed and an account would be called for. But most North American universities have assessment offices and take assessment much more seriously (in many ways) than do their British counterparts. One consequence is that they actually have to hand detailed data about student performance that can be centrally analysed in all sorts of ways. In the UK, data tend to be held locally, with university administrations often holding only the most bland of data, of little value even if they were disposed to ask some awkward questions about the actual operation of the validated assessment system.

The point of this account of assessment practices at Arcadia is not to claim that there is anything particularly unusual there, except only that unusual interest seems to have been taken in finding out what is happening within the assessment system. The point is that existing assessment systems, such as Arcadia's, advantage some learners and disadvantage others; reward some forms of achievement and not others; and seem to do so in eclectic ways, without evidence of any unifying rhyme or reason. Proponents of new approaches to assessment should have to account for the values that are being promoted by their innovations, for by promoting *these* values they are narrowing the scope within which *those* values can operate. However, it is not a contest between the innovators' values and an existing, morally-neutral system. Rather, the present system is as open to moral objection as any other and perhaps, because of its seemingly ramshackle nature, it is more open to objection.

OVERVIEW OF THE BOOK

As Madeleine Atkins says in Chapter 1, these arguments hinge on what we expect higher education to do:

> Inescapably, the issues are about what students are learning and who is going to define it. And the answers cannot just be in terms of rather low-level generic skills and competences if higher education is to justify its costly existence in the twenty-first century. There also has to be a debate about the knowledge and understanding that one should expect a student to gain from an undergraduate programme, unfashionable though such a debate has been for some time....

It is a truism that effective assessment depends upon having a view of what it is that we are trying to do in a programme, hence of what it is that we ought to assess. What is distinctive and important about this chapter is the crisp review of four competing claims about the purpose of higher education. In the process of reviewing them, she notices a number of awkward questions about the development of expertise, the notion of the reflective practitioner and the validity of the concept of 'general transferable skills', for example. The general thrust is that too many claims about the contribution of higher education have been founded on rhetoric rather than on careful analysis with attention to appropriate evidence. In this respect her chapter might usefully be read alongside Barnett's (1994) stimulating views on the nature of higher education.

Arguing that higher education should probably be aiming to provide a general educational experience of intrinsic worth in its own right and to

prepare students for general employment, she recognizes that this position has considerable implications for teaching methods, assessment issues and staff training. These, though, are second-order matters. First, we need greater clarity about what we are trying to do through higher education and that clarity ought to be a product of the application of our analytical and critical powers, not a result of their absence. One conclusion that might be drawn from the case of the University of Arcadia is that clarity of purpose is lacking, which can be discerned in the assessment arrangements, which seem to have been hardly touched by the Enterprise in Higher Education Project money that has flowed through the university.

David Boud concentrates on what assessment systems do to learners, which is a recurring theme in this book (see Chapters 4 and 6, for example). He uses the concept of consequential validity ('the effect of the test or other form of assessment on learning and other educational matters') in examining ways in which assessment systems help or hinder the types of learning that we say we wish to encourage. 'Assessment', he says, 'is the most significant prompt for learning'; it 'acts as a mechanism to control students that is far more pervasive and insidious than most staff would be prepared to acknowledge'. Unfortunately, academic staff are not as sensitive as they might be to the way assessment seems when seen through student eyes. 'Even successful, able and committed students ... have been hurt by their experiences of assessment, time and time again'. Nor is it enough for any one academic to try to mitigate such effects, for,

> in any given month they [students] may have to complete ten assessment tasks, in another month only one. The ways in which they approach each of these will be influenced by the others... Very little attention has been given to the compounding effects of assessment.

Looked at in this way, assessment reform is not simply something for enthusiastic academics to undertake, but is a matter for departmental and institutional action, which also follows from the arguments developed in Chapter 1. Illustrations of university-wide approaches to assessment reform are to be found in Chapters 7–9.

Yet, action by individual academics is important. To repeat a cliché, effective change is simultaneously 'bottom-up' and 'top-down', a combination of tinkering and radical overhaul. Take the language of assessment as an example. Boud observes that the language of assessment is often excessively judgemental; 'it has the final say. It classifies without recourse to reconsideration or further data. And it does not allow for further possibilities'. 'We judge too much and too powerfully, not realizing the extent to which students experience our power over them'.

While raising awareness of this problem is best done on an institution- and department-wide basis, individual academics can contribute a lot by reappraising their practice and experimenting with better ways of giving feedback to students.

Dary Erwin is the Director of the Office of Assessment at James Madison University, Virginia. Such offices are to be found in most American universities, signalling a university commitment to taking seriously the assessment of student learning and of the learning experience. Like Boud, he is interested in the relationship between assessment and learning but he approaches the issue from an account-ability and course improvement perspective. There is, he observes, a crisis in assessment, so that,

> grade inflation, awarding grades based on effort and not perfor-mance, uneven standards among instructors or among institutions, and a lack of understanding or agreement about education itself have led to a lack of credible measures.

Echoing Madeleine Atkins, he insists that 'often the lack of clarity in objectives, purposes, competences, or whatever term one wishes to use, causes confusion about what is in the curriculum'.

Despite these internal problems, it is imperative for universities to be able to produce valid and reliable assessment data to demonstrate their value at a time when it is increasingly difficult to secure state resources, and in order to undertake well-informed programme development so as to enhance the student learning experience. His conclusion is that,

> to ignore calls for accountability is to encourage people external to higher education to establish their standards rather than ours. Our response is more critical now than ever. And the credibility of our response depends to a large degree on the adequacy of the assess-ment process we have in place.

It is a conclusion that cannot be ignored anywhere in the anglophone world.

In Chapter 4, Phil Race offers a variety of checklists to help academics, managers, funding councils, quality auditors and policy-makers to scrutinize existing assessment practice in the areas for which they have responsibility. These lists represent the pooled wisdom of delegates to the SEDA conference and as such carry the authority of massed expertise. This authority is enhanced by the data presented in the first part of the chapter, which go a long way to affirm points made in Chapter 2. These delegates, academically and professionally successful people (one trusts),

have nightmares about being assessed; have more negative feelings about exams and coursework than they do positive ones, even though they must, as academics, know the case *for* assessment; thought that there is little learning pay-off from examinations; and had concerns about both examinations and coursework assessment. Comparison of these data with the four-fold model of learning which Race advances suggests that assessment practices are, at best, not helping the learning process and, at worst, injurious to it.

Race has also contributed to Chapter 5, in which Chris Rust and Sally Brown briskly discuss a number of questions that might be asked about assessment if we were to take heed of the points made in the previous four chapters and critically appraise its fitness for purpose. They too emphasize the importance of clarity of purpose but they make a point of saying that this must be a *shared* clarity. Learners need to understand the criteria by which their work will be judged every bit as much as do the tutors – arguably more so. After all, assessment is 'an engine for learning'.

This process can be helped by peer- and self-assessment, which are assessment modes that are also valuable on other grounds, such as efficiency. Boud is recognized as a leading proponent of self-assessment as a way of fostering learner autonomy and Brown, Race and Rust follow other contributors to this book in endorsing this position. It seems as though self- and peer- assessment are ideas whose time has now come.

The importance of good feedback is also stressed (and this is a theme that is revisited in Chapter 12), and Records of Achievement, or profiles, are identified as a way in which the meanings of tutors' judgements and of students' perspectives may be brought together and, through negotiation, worked into plans for future personal development within a programme of study. As with Boud, these writers are concerned that through assessment 'the tutor can make learners feel powerless', given 'the seemingly arbitrary nature of tutor assessment' and consider ways in which assessment processes might begin to 'empower' the learner.

This is also a theme of Chapter 6 by Bob Farmer and Diana Eastcott, who adopt Race's model of learning in their review of ways in which assessment might be used to enhance the quality of the student learning experience. As a drawing-together chapter, this endorses much that has gone before but adds to that material examples of practices in use, places greater emphasis on portfolios as a method of formative assessment and draws on Kolb's learning cycle in considering ways of helping students to reflect upon their own learning. Their conclusion, that '"How am I going to be tested?" is often at the very heart of students' approaches to and feelings about learning', stands as a summary of the main thrust of these first six chapters.

The same themes are woven through the next six chapters, although

these chapters are more in the nature of reports of work that has been or is being done on assessment, rather than being analyses of the place and attributes of assessment systems in general.

Ivan Moore describes a rare example of a university-wide approach to assessment reform prompted, interestingly enough, by a feeling that too few first-class degrees were being awarded. His study offers an interesting complement to the case developed in a recent SEDA paper on staff development and university-wide change (Knight, 1994), showing how assessment, teaching and learning are intertwined and that action is necessary at all levels of a university. This makes for a demanding job for the staff developer, although, as Moore points out, this particular project offers greater leverage on educational development than would many others. He presents estimates of the time-cost of the programme that suggest that it will prove to be a very efficient way of provoking significant thinking about the university's teaching and of supporting changes, especially in assessment practices. The following long-term benefits are anticipated:

- changes in course design

- long-term change

- staff will be better informed to continue to improve their assessment practices.

At the University of Plymouth, change agents were used to stimulate assessment reform. An interesting feature, described by Hazel Fullerton in Chapter 8, was an emphasis on using visual media as a way of sharing the development of thinking about assessment reform and then sharing the conclusions with others. This process is succinctly described and five examples are given of posters produced at Plymouth. While the spur to this development was increasing student numbers, the combination of a 'bottom-up' strategy and support from senior management appears to have worked so well that seven developments in assessment practice are now embedded within the university.

Kate Day and Dai Hounsell describe their programme to help part-time graduate tutors to assess better. This, as they remark, is an issue of considerable importance, since some universities are divesting more and more teaching to untrained graduate assistants so that full-time staff can concentrate on the research selectivity exercise. In Chapter 9 they discuss the problems which face staff and educational development personnel when they try to help such a diverse group of people to develop their skill at assessing student learning. Prudently, they conclude that only so much can be done on a university-wide basis, while still insisting that,

it is difficult to see how from the viewpoints of quality and accountability universities can avoid setting firm and explicit institution-wide boundaries ... and thus provide more robust policy and practical frameworks within which course and departmental tailoring can be constructively accommodated.

The interplay of assessment and learning is explored in Chapter 10 by Phyllis Creme. She argues that one result of changes at the margins of course assessment requirements has been that 'group discussions have become a more explicit and structured part of the course', with evidence of students taking greater responsibility for their learning. However, the issue of control, raised in Chapters 2 and 4, was not resolved, for while 'students were being expected to take more responsibility than previously for the seminar process', assessment was still something done to them, so that they 'can hardly be expected to be open – or adult – in their attitudes to knowledge if they are not, in the end, allowed to take responsibility for assessing it'.

Similar themes are explored by Irene Harris in Chapter 11, who investigates assessment issues attaching to the use of learning contracts. While learning contracts promise a more equal relationship between learners and tutors, the promise works out quite differently on the three courses that she studied: 'the actual [assessment] situation for individual learners may be very different to the outward experience'. 'Letting go', she added, quoting Tomkins and McGraw (1988, p. 177), 'is sometimes the greatest challenge for the teacher'. So, if learning contracts are to make good their promise, both teachers and students will need to learn new skills.

One of those skills, foreshadowed by Chapters 10 and 11, is that of peer-assessment, which readily shades into self-assessment. Yet peer-assessment causes much concern in some quarters. In Chapter 12, Nancy Falchikov describes her approach to using peer-assessment to improve the quality and speed of feedback to students. Her conclusions were that students' marks were closely aligned with tutors' marks; that motivation benefited; that students were inclined to use tentative language in their comments (Chapter 2, q.v.); and that students were required to use reflection and engage in critical analysis. However, 'it sometimes appeared that only when seminar delivery was good or adequate did students look beyond it to issues of structure and content'. This might suggest that unless students are progressively helped to become more perceptive in their analyses, then peer feedback marking and associated assessment methods will prove to be useful only for assessing the surface features of presentations, which would severely limit the appeal of these widely promoted approaches (see Chapter 5, for example).

This is a salutary reminder that many ideas that are recommended as ways of improving assessment practices have face validity (they look sensible) but they are not supported by the range of systematic and well-designed studies that would allow greater confidence to be placed in them. Perhaps this is inevitable in these early days of reforming assessment in the interests of better learning. Writing of research into school examining, Wood (1991, p. 245) wondered whether 'it is only when a change is announced, preparations for it are made and awareness is raised that serious research becomes possible'. Moreover, 'if it is clear that innovation is not driven by research, and that the reverse may frequently be true, there remain some pointed questions about innovation that beg to be explored' (p. 248). Indeed. As I write this introduction I am conscious on the one hand of the enormous interest in and creativity about assessment and, on the other, by how very little we know, even in the University of Arcadia (but perhaps not at James Madison University), about the effects of assessment on learners, as well as on learning, teaching and teachers.

The National Council for Vocational Qualifications is an example of a political innovation that called forth an enormous amount of research to shape, underpin and refine National Vocational Qualifications (NVQs) and General National Vocational Qualifications (GNVQs). This work has enormous significance for higher education, as Romla Hadrill explains in the final chapter. Acknowledging that there are ideological and technical disputes surrounding the NCVQ's work (see also Barnett, 1994; Eraut, 1994b; Hodkinson and Issitt, 1994; Hyland, 1994; Tomlinson and Saunders, 1995), she none the less sets out the thrust of this work. Higher education would be unwise to assume that NVQs and GNVQs can be ignored. First, the approach to assessment of specifying learning outcomes and then proceeding to assess learners' competence against them will not go away. Moreover, the development of various accreditation and credit accumulation schemes implies that some universities will be interested in recognizing competence that has not been developed through university courses, which is a basic NCVQ principle. Lastly, NVQs have already been developed to levels of achievement equivalent to university work and GNVQs are to be developed to these levels. Not only may these developments have a direct influence on higher education but it is a small step to the idea of a core curriculum for higher education based around the development of certain key competences.

This takes us back to Madeleine Atkins's fundamental questions. What is higher education for? So what are we to assess, given that assessment is a moral activity? To these another two, which have been an undertone throughout this introduction, might be added: What data are available about the effects of different forms of assessment upon different groups of students? And are these data good enough?

References

Abercrombie, MLJ (1989) The Anatomy of Judgement: An investigation into the processes of perception and reasoning, London: Free Association Books.

Atkins, MJ , Beattie,J. and Dockrell, WB (1993) Assessment Issues in Higher Education, Sheffield: Employment Department.

Barnett, R. (1994) The Limits of Competence, Buckingham: Open University Press/SRHE.

Baume, D. and Brown, S. (eds) (1992) Learning Contracts, volume two, Some practical examples, Birmingham: Standing Conference on Educational Development.

Becher, T. (1989) Academic Tribes and Territories, Buckingham: SRHE and OpenUniversity Press.

Benne, KD , Bradford, LP and Lippett, R. (1964) 'The laboratory method', in Bradford, L. , Gibb JR and Benne, KD (eds) T-Group Theory and Laboratory Method, New York: John Wiley and Sons.

Bennett, RE and Ward, WC (eds) (1993) Construction Versus Choice in CognitiveMeasurement: Issues in constructed response, performance testing, and portfolio assessment, Hillsdale, NJ: Lawrence Erlbaum Associates.

Berk, RA (1986) 'A consumer's guide to setting performance standards on criterion-referenced tests', Review of Educational Research, 56, 137–172.

Bloom, BS (1956) Taxonomy of Educational Objectives, volume 1, The cognitive domain, London: Longman.

Boud, D. (ed.) (1988a) Developing Student Autonomy in learning 2nd edn, London: Kogan Page.

Boud, D. (1988b) 'Moving towards autonomy', in Boud, D. (ed.) Developing Student Autonomy in Learning 2nd edn, London: Kogan Page.

Boud, D. (1991) Implementing Student Self Assessment, HERDSA Green Guide, 2nd edn, Sydney: Higher Education Research and Development Society of Australasia.

Boud, D. and Feletti, G. (1992) The Challenge of Problem Based Learning New York: San Martin's Press.

Brown, G. and Pendlebury, M. (1992) Assessing Active Learning Sheffield,: CVCP Universities Staff Development and Training Unit.

Brown, S. and Baume, D. (eds) (1992) Learning Contracts, volume one: A theoretical perspective, Birmingham: Standing Conference on Educational Development.

Brown, S. and Knight, P. (1994) Assessing Learners in Higher Education, London: Kogan Page.

Brown, S. , Rust, C. and Gibbs, G. (1994) Strategies for Diversifying Assessment, Oxford: Rewley Press.

Bruner, J. (1970) 'Some theorems on instruction', in Stones, E. (ed.) Readings in Educational Psychology, London: Methuen.

Bruner J. (1992) 'Another look at New Look 1', American Psychologist, 47, 6, 780–783.

Channel Four (1993) 'All Our Futures – Britain's Education Revolution', London: Channel Four TV.

Dahlgren, L-O (1984) 'Outcomes of learning', in Marton, F. , Hounsell, D. and Entwistle, N. (eds) The Experience of Learning Edinburgh: Scottish Academic Press.

Ebel, RL and Frisbie, DA (1986) Essentials of Educational Measurement, Engelwood Cliffs, NJ: Prentice Hall.

Ecclestone, K. (1994) 'Democratic values and purposes: the overlooked challenge of competence', fournal of Educational Studies, 20, 2.

Eisner, EW (1993) 'Reshaping assessment in education: some criteria in search of practice', Journal of Curriculum Studies, 25, 3, 219–233.

Elton, L. et al (1994) Staff Development in Relation to Research, Sheffield: CVCP Universities' Staff Development Unit.

English, L. (1992) 'Children's use of domain-specific knowledge and domain-general strategies in novel problem solving', British Journal of Educational Psychology, 62, 203–216.

Entwistle, NJ (1992) The Impact of Teaching on Learning Outcomes in Higher Education, Sheffield: Committee of Vice-Chancellors and Principals (Staff Development Unit).

Entwistle, NJ and Ramsden, P. (1983) Understanding Student Learning Beckenham: Croom Helm.

Entwistle, NJ and Entwistle, A. (1991) 'Contrasting forms of understanding for degree examinations: the student experience and its implications', Higher Education, 22, 205–227.

Eraut, M. (1994a) Ethics in Occupational Standards NVQs and SVQs, Sheffield: Employment Department, Learning Methods Branch.

Eraut, M. (1994b) Developing Professional Knowledge and Competence, London: Falmer Press.

Eraut, M. and Cole, G. (1993) 'Assessment of competence in higher level occupations', Compendium No 3. Competence and Assessment, Sheffield: Employment Department, Learning Methods Branch.

Erwin, T. (1983) 'The scale of intellectual development: measuring Perry's scheme', Journal of College Student Personnel 24, 6–12.

Erwin, TD (1991) Assessing Student Learning and Development, San Francisco, CA: Jossey-Bass.

Falchikov, N. (1986) 'Product comparisons and process benefits of collaborative self and peer group assessments', Assessment and Evaluation in Higher Education, 11, 2, 146–166.

Falchikov, N. (1988) 'Self and peer assessment of a group project designed to promote the skills of capability', Programmed Learning and Educational Technology, 25, 4, 327–339.

Falchikov, N. (submitted for publication) 'Peer Feedback Marking: developing peer assessment'.

Forehand, LS , Vann, WF and Shugars, DA (1982) 'Student self-evaluation in pre-clinical restorative dentistry', Journal of Dental Education, 46, 4, 221–226.

Forster, F. and Hounsell, D. (1994) Tutoring. A handbook for postgraduate and other part-time tutors, Edinburgh: Centre for Teaching, Learning and Assessment.

Gibbs, G. (1981) Teaching Students to Learn, Buckingham: Open University Press.

Gibbs, G. (1991) 'Reflection and self assessment for new lecturers', in Brown, S. and Dove, P. (eds) Self and Peer Assessment, Birmingham: Staff and Educational Development Association.

Gibbs, G. (1992) Improving the Quality of Student Learning Bristol: Technical and Educational Services.

Gibbs, G. (in press). 'Changing lecturers' conceptions of teaching and learning through action research', in Brew, A. (ed.) Directions in Staff Development, Buckingham: SRHE/Open University Press.

Gibbs, G. , Habeshaw, S. and Habeshaw, T. (1988) 53 Interesting Ways of Assessing Your Students, 2nd edn, Bristol: Technical and Educational Services.

Gonczi, A. (1994) 'Competency based assessment in the professions in Australia', Assessment in Education, 1, 1, 27–44.

Gray, TGF (1987) 'An exercise in improving the potential of exams for learning', European Journal of Engineering Education, 12, 4, 311–323.

Hager, P. , Gonczi, A. and Athanasou, J. (1994) 'General issues about assessment of competence', Assessment and Evaluation in Higher Education, 19, 1, 3–16.

Harvey, L. (1993) 'An integrated approach to student assessment', Warwick: Paper presented to the Measure for Measure, Act II Conference, 6–8 September.

Hayes, E. and Colin, SAJ (eds) (1994) Confronting Sexism and Racism, New Directions for Adult and Continuing Education, 61, San Francisco, CA: Jossey-Bass.

Heron, J. (1981) 'Assessment revisited', in Boud, D. (ed.) Developing Student Autonomy in Learning London: Kogan Page.

Heywood, J. (1989) Assessment in Higher Education, 2nd edn, Chichester: john Wiley and Sons.

Hodkinson, P. and Issitt, M. (1994) The Challenge of Competence, London: Cassell.

Hounsell, D. and Murray, R. (1992) Essay Writing for Active Learning Sheffield: CVCP Universities' Staff Development Unit.

Hustler, D. , Peckett, J. and Whiteley, M. (1993) Learning Contracts and Initial Professional Development: A case study from initial teacher education, Manchester: Didsbury School of Education.

Hyland, T. (1994) Competence, Education and NVQs, London: Cassell.

Jacobs, A. (1974) 'The use of feedback in groups', in Jacobs, A. and Spradlin, WW (eds) The Group as Agent of Change, New York: Behavioral Publications.

Jessup, G. (1991) Outcomes – NVQs and the emerging model of education and training London: Falmer Press.

Jordan, TE (1989) Measurement and Evaluation in Higher Education, London: Falmer Press.

Knight, P. (ed.) (1994) University-wide Change, Staff and Curriculum Development, Birmingham: Staff and Educational Development Association.

Knight, P. (1995) Records of Achievement in Further and Higher Education, Lancaster: Framework Press.

Knottenbelt, M. and Fiddes, N. (1994) Part-Time Tutoring. A survey of teaching by postgraduate and other part-time tutors at the University of Edinburgh, Edinburgh: Centre for Teaching, Learning and Assessment.

Knowles, M. (1986) Using Learning Contracts, San Francisco, CA: Jossey Bass.

Kohn, A. (1993) Punished by Rewards, Boston, Mass.: Hough ton Mifflin.

Kolb, D. (1984) Experiential Learning Englewood Cliffs, NJ: Prentice Hall.

Korman, M. and Stubblefield, RL (1971) 'Medical school evaluation and internship performance', Journal of Medical Education, 46, 670–673.

Laurillard, DM (1984) 'Learning from problem solving', in Marton, F. , Hounsell, D. and Entwistle, N. (eds), The Experience of Learning Edinburgh: Scottish Academic Press.

Linn, BS , Arostegui, M. and Zeppa, R. (1975) 'Performance rating scale for peer and self assessments', British Journal of Medical Education, 9, 98–101.

Linn, RL , Baker, EL and Dunbar, SB (1991) 'Complex, performance-based assessment: expectations and validation criteria', Educational Researcher, 20, 8, 15–21.

Long, DG (1990) Learner Managed Learning London: Kogan Page.

Luke, C. and Gore, J. (eds) (1992) Feminisms and Critical Pedagogy, London: Routledge.

MacDonald, RB (1991) 'Developmental students' processing of teacher feedback in composition instruction', Review of Research in Developmental Education, 8, 5.

MacFarlane Report (1992) Teaching and Learning in an Expanding Higher Education System, Edinburgh: Committee of Scottish University Principals.

Marton, F. and Saljo, R. (1976) 'On qualitative differences in learning – outcomes as a function of the learner's conception of the task', British Journal of Educational Psychology, 46, 115–127.

Marton, F. and Saljo, R. (1984) 'Approaches to learning', in Marton, F. , Hounsell, D. and Entwisde, N. (eds) The Experience of learning Edinburgh: Scottish Academic Press.

Mast, TA and Bethart, H. (1978) 'Evaluation of clinical dental procedures by senior dental students', Journal of Dental Education, 42, 4, 196–197.

Mathews, BP (1994) 'Assessing individual contributions: experience of peer evaluation in major group projects', British Journal of Educational Technology, 5, 1, 19–28.

Messick, S. (1989) 'Validity', in Linn, RL (ed.) Educational Measurement, 3rd edn, New York: Macmillan.

Miller, CML and Parlett, M. (1974) Up to the Mark: A study of the examination game, Guildford: SRHE.

Mitchell, L. and Bartram, D. (1994) The Place of Knowledge and Understanding in the Development of National Vocational Qualifications and Scottish Vocational Qualifications, Sheffield: Employment Department.

Morton, J.B. and Macbeth, W.A.A.G. (1977) 'Correlations between staff, peer, and self assessments of fourth-year students in surgery', Medical Education, 11, 3, 167–170.

National Center for Education Statistics (1992) National Assessment of College Student Learning: Issues and concerns, Washington, DC: US Government Printing Office.

Osterlind, SJ (1989) Constructing Test Items, Norwell, MA: Kluwer Academic Publishers.

Otter, S. (1992) Learning Outcomes In Higher Education, London: Department for Education.

Parry, M. (1994) 'Renaissance virtue gets the soft sell', The Times Higher Education Supplement, 22 April, 22.

Perry, WG (1970) Forms of Intellectual and Moral Development in the College Years: A scheme, New York: Holt, Rinehart and Winston.

Pettit, P. (ed.) (1993) Consequentialism, Aldershot: Dartmouth Publishing Co.

Pettman, JJ (1991) 'Towards a (personal) politics of location', Studies in Continuing Education, 13, 2, 153—166.

Polytechnics and Colleges Funding Council/Council for National Academic Awards (1990) The Measurement of Value Added in Higher Education, London: CNAA.

Race, P. (1992) 'Not a learning contract', in Brown and Baume , op cit

Race, P. (1994a) The Open Learning Handbook, 2nd edn, London: Kogan Page.

Race, P. (1994b) Never Mind the Teaching, Feel the learning, Birmingham: Staff and Educational Development Association.

Ramsden, P. (1986) 'Students and quality', in Moodie, GC (ed.) Standards and Criteria in Higher Education, Guildford: SRHE and NFER-Nelson.

Ramsden, P. (1987) 'Improving teaching and learning in higher education: the case for a relational perspective', Studies in Higher Education, 12, 275–286.

Ramsden, P. (1988) 'Studying learning: improving teaching', in Ramsden, P. (ed.) Improving Learning: New perspectives, London: Kogan Page.

Rogers, CR (1983) Freedom to Learn for the 80's, Columbus, Ohio: Merrill.

Roid, GH and Haladyna, TM (1982) A Technology for Test-Item Writing New York: Academic Press.

Rorty, R. (1989) Contingency, Irony and Solidarity, Cambridge: Cambridge University Press.

Scriven, M. (1967) 'The methodology of evaluation', in Tyler, RW et al (eds) Perspectives of Curriculum Evaluation, Chicago: Rand McNally.

Stephenson, J. and Laycock, M. (eds) (1993) Using Learning Contracts in Higher Education, London: Kogan Page.

Stones, E. (1970) 'Strategy and tactics in programmed instruction', in Stones, E. (ed.) Readings in Educational Psychology, London: Methuen.

Terenzini, P. (1993). 'Cross-national themes in the assessment of quality in higher education', Assessment Update, 5, 1–14.

Thomas, K. (1990) Gender and Subject in Higher Education, Buckingham: SRHE/Open University Press.

Tomkins, C. and McGraw, M-J (1988) 'The negotiated learning contract', in Boud, D. (1988a), *op cit*

Tomlinson, P. and Saunders, S. (1995) 'The current possibilities for competence profiling in teacher education', in Edwards, A. and Knight, P. (eds) The Assessment of Competence in Higher Education, London: Kogan Page.

Turner, K. (1993) 'An investigation of how students respond to feedback on coursework', Warwick: Paper presented to the 'Measure for Measure Act II' Conference, September.

Wesman, AG (1972) 'Writing the test item', in Thorndike, RL (ed.) Educational Measurement, Washington, DC: American Council on Education.

Wiggins, G. (1989) 'A true test: toward more authentic and equitable assessment', Phi Delta Kappa, 71, 9, 703–713.

Winter, R. (1995) 'The assessment of professional competences: the importance of general criteria', in Edwards, A. and Knight, P. (eds) The Assessment of Competence in Higher Education, London: Kogan Page.

Wood, R. (1991) Assessment and Testing Cambridge: Cambridge University Press.