Assessment for Learning *in* Higher Education

SEDA

EDITED BY
PETER KNIGHT

Staff and Educational Development Series

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, networking, 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

LONDON AND NEW YORK

Published in association with the
Staff and Educational Development Association

First published in 1995
By RoutledgeFalmer
Reprinted in 1998

Reprinted 2004
By RoutledgeFalmer
2 Park Lane
Milton Park
Abingdon
Oxon
OX14 4RN

Transferred to Digital Printing 2005

RoutledgeFalmer is an imprint of the Taylor & Francis Group

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms of licences issued by the Copyright Licensing Agency.

© Peter Knight and named contributors, 1995

British Library Cataloguing in Publication Data

A CIP record for this book is available from the British Library.

ISBN 0 7494 1532 0

Typeset by DP Photosetting, Aylesbury, Bucks

Contents

Notes on contributors 7

Preface 11

Introduction 13
Peter Knight

1. What Should We be Assessing? 25
Madeleine Atkins

2. Assessment and Learning: Contradictory or Complementary? 35
David Boud

3. Attending to Assessment: A Process for Faculty 49
T Dary Erwin

4. What has Assessment Done for Us – and to Us? 61
Phil Race

5. Using and Experiencing Assessment 75
Sally Brown, Phil Race and Chris Rust

6. Making Assessment a Positive Experience 87
Bob Farmer and Diana Eastcott

7. Staff and Educational Development for Assessment Reform: A Case Study 95
Ivan Moore

8. Embedding Alternative Approaches in Assessment 111
Hazel Fullerton

9. When Tutors Assess: Who Can Help and How? 125
Kate Day and Dai Hounsell

10. Assessing 'Seminar' Work: Students as Teachers 137
Phyllis Creme

11. Learning Contracts – Related Assessment Issues 147
Irene Harris

12. Improving Feedback To and From Students 157
Nancy Falchikov

13. The NCVQ Model of Assessment at Higher Levels 167
Romla Hadrill

References 181

Index 187

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.*

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 accountability and course improvement perspective. There is, he observes, a crisis in assessment, so that,

> grade inflation, awarding grades based on effort and not performance, 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 assessment 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),