

# ASSESSMENT IN HIGHER EDUCATION

---

SECOND EDITION

---

John Heywood

# *Assessment in Higher Education*

*Second Edition*

JOHN HEYWOOD

*Trinity College, The University of Dublin*

JOHN WILEY & SONS

Chichester · New York · Brisbane · Toronto · Singapore

Copyright © 1977, 1989 by John Wiley & Sons Ltd.

All rights reserved.

No part of this book may be reproduced by any means,  
or transmitted, or translated into a machine language  
without the written permission of the publisher

***Library of Congress Cataloging-in-Publication Data:***

Heywood, John, 1930–  
Assessment in higher education / John Heywood.—2nd ed.  
p. cm.

ISBN 0 471 92032 0

1. Universities and colleges—Examinations. 2. College teaching—Evaluation.  
3. Education, Higher—Evaluation. 4. Education, Higher—Aims and objectives  
I. Title.

LB2366.H49 1988

378'.166—dc19

88-14423  
CIP

***British Library Cataloguing in Publication Data:***

Heywood, John, 1930–  
Assessment in higher education.—2nd ed.  
1. Higher education institutions. Students. Academic achievement.  
Assessment. I. Title  
378'.1664

ISBN 0 471 92032 0

Filmset by Bath Typesetting Ltd, Bath, Avon

Printed and Bound in Great Britain by Anchor Press Ltd, Tiptree, Essex

# *Assessment in Higher Education*

*For  
Pauline*

## *Preface*

When the first edition of this book was published in 1977, while the problems of testing and assessment in respect of reliability and validity were well understood, assessment of students remained an afterthought of the educational process. It was felt that tests and examinations had a backwash effect on student learning, but exactly how that effect could be used to enhance it was little investigated. My purpose in the first edition was to explore the relationship between assessment, instruction and learning, and that continues to be the purpose of this edition.

I argued that assessment had to be seen as an integral part of the learning process. The objectives of assessment, learning and instruction were the same. Thus the instructional procedure used by a teacher in higher education had to be selected so as to achieve the objectives of instruction, and at the same time the assessment techniques had also to be a valid test of those same objectives. The consequences of this view for class and course design were profound, for the time required for the completion of the units of instruction and their assessment determined what could be accomplished in a course. In consequence, the task of selecting a few significant aims and objectives was of great importance in the design of assessment and instruction. This task, which, following E. J. Furst, I called 'screening', depended on the teacher in higher education evolving an epistemological basis for his work together with a defensible theory of learning as a foundation for the attainment of those aims and objectives thought to be significant. The many sources of aims and objectives were discussed and the activity of instruction and assessment design was shown to be complex. In the closing chapters it was argued that there was an obligation on teachers to evaluate their teaching and that the model of illuminative evaluation provided a way forward.

Much has happened since then. A great deal has been written on forms of assessment other than written tests, on anxiety and tension resulting from testing and instruction, objectives, teaching and learning styles, and the ways in which examinations and tests influence learning, which all has a bearing on the activity of screening. Nevertheless, this work has been done by a small band of individuals and not on an institution-wide basis.

At the same time, the managers of higher education are calling for institutional approaches to the problems of teaching and learning. In Britain the recommendations of a committee of the Committee of Vice-Chancellors and Principals chaired by Professor P. A. Reynolds on external examining have

been acted on. These call for the assessment of staff (appraisal), the assessment of teachers by students (student evaluation), and the review of departments. In at least one university ideas gained from the accreditation techniques used in the United States have been seriously considered.

The effectiveness with which such judgements can be made depends on the willingness of institutions and the individuals within them to be clear about the 'significant' aims and objectives which they wish to obtain. Thus the model of assessment discussed in this book is also that for institutional evaluation.

In the first edition it was argued that no single technique of assessment could be expected to evaluate even a few significant objectives. For this reason, a multiple-objective approach to teaching and learning had to be taken to the design of the curriculum, its assessment and instruction. This point is made in the recent report of the Governors of the United States on the assessment of quality in higher education. They believe that the assessment of institutions of higher education and the individuals within them should be multi-strategy in approach.

Ultimately the success of this newfound movement in assessment and teaching in higher education will depend on the seriousness with which institutions take the remit and the ways they find to encourage teachers not only to be good research workers but good researchers of their own classroom practice. For this, the teachers themselves will need to become practised in the reflective thinking about the art of their teaching or, as it is commonly called today, self-assessment. This book is intended to provide a basis for practice and reflection on assessment in higher education.

## *Acknowledgements*

I am especially grateful to Dr Georgine Loacker of Alverno College, Milwaukee, for her continuing support of my work over the years. I am also grateful to the President of Alverno and her colleagues for the welcome which they extended to my wife and myself when we visited the College. Their introductions led to encouragement and help from Dr Russel Edgerton, President of the American Association for Higher Education and his colleagues. Dr Clifford Adelman of the Office of Educational Research and Improvement, US Department of Education, provided a continuous flow of 'fugitive literature' as well as introductions to other persons working in the field.

My friends Dr James Freeman of the University of Manchester and Dr Seamus McGuinness of the University of Dublin, who helped me through the first edition, once again came to my support. I am grateful to them for their many helpful suggestions, as I am also to the publishers' unknown referee.

Finally, I would like to thank the University of Dublin, Trinity College, for the grant which enabled me to visit the United States and the Dean of the Faculty of Arts, Dr Patrick Kelly, for all his help.



# Contents

<b>Preface</b>	<b>xi</b>
----------------	-----------

<b>Acknowledgements</b>	<b>xiii</b>
-------------------------	-------------

<b>1 Assessment in Higher Education: An Introduction</b>	<b>1</b>
--	----------

Attending to assessment in higher education: an historical note—Selection for higher education and the system of education—Meanings of assessment, evaluation and appraisal in higher education—Assessment by students of teacher performance—The idea of multiple-objective and multiple-strategy assessment

<b>2 En Route to a Degree</b>	<b>27</b>
-------------------------------	-----------

Introduction—Systems approaches to the study of the impact of college on the student pursuit of grades—Society, education and grades—Grades, professionalism and the system of education—Admissions policies—The meaning that grades have for employers

<b>3 Grades and Grading</b>	<b>45</b>
-----------------------------	-----------

Introduction—The pervasive influence of examinations and tests—Differential grading between subjects—Grading: absolute or relative?—The perceptions of examiners and the combination of their marks—Differential subject performance: the problem of comparability—Grading practices and problems of selection—External examining

<b>4 Temperament and Testing</b>	<b>76</b>
----------------------------------	-----------

Introduction—Test anxiety—Temperament and performance—Examination strain (testing and mental health)—Coursework, motivation and stress—Course anxiety—The short- and long-term goals of students and achievement—Implications for assessment, course structure and the role of the teacher

<b>5 Defining the Task of Assessment: Implications for Teaching</b>	<b>105</b>
---	------------

Introduction—The importance of aims—A problem of terminology: meanings of the term 'objective'—The taxonomy of educational objectives—To declare or not to declare objectives—Intended versus unintended outcomes—Critic and connoisseur: the art of self-assessment in teaching

## **6 Finding Objectives for Learning and Competency 123**

Introduction—Critical thinking: a philosophical perspective—Critical thinking: a psychological perspective—Using the *Taxonomy* in evaluation and assessment research—Goal clarification using the Delphi technique and other methods of group problem-solving—A critical thinking project—Using the *Taxonomy* in job analysis—Deriving education and training programmes from job analyses—The value-added dimension of assessment—Work and proficiency: the Critical Incident Technique as a means of deriving objectives—Competency-based assessment and learning—Conclusions

## **7 Concept Learning and Curriculum Objectives 157**

Introduction: towards the clarification and assembly of goals—Learning theories—Introducing chapters 7–9—The importance of concepts and principles: instruction and the curriculum—The entering characteristics of students—Alternative instructional strategies—Curriculum structures for concept learning

## **8 Perception, Learning and Studying 177**

Introduction—Learning and experience—Interpreting the outer world—Meaningful learning and the acquisition of concepts—Value dispositions and cognitive dissonance—Meeting the goals of perceptual learning in the curriculum—Spatial ability—Field-dependence and field-independence—Convergent and divergent thinking styles—Kolb's learning styles—Study skills: learning how to learn—Depth and surface approaches to learning

## **9 Learning and the Role of the Teacher 203**

Introduction—Personal and intellectual development—Piaget's theory of cognitive development—Kohlberg's theory of moral development—Perry's theory of development in higher education—Loneragan's theory of human knowing—A general criticism of theories of development—The language of subjects: the implication of development theory for the curriculum—Learning, personality and the teacher—Assessment, instruction and the role of the teacher in higher education

## **10 Written Examinations and Assignments (content, structure and marking) 225**

Introduction—Literacy and numeracy—The marking of essays and problem-solving examinations—An approach to the marking of large numbers of

scripts—Question design and the design of essay examinations—Question spotting: perceptions of assessment and performance—Student perceptions of essay writing—Feedback—Short-answer questions—Open-book examinations—The timing of examinations: prior notice (seen) questions

## **11 Practicals, Projects, Profiles and Self-assessment 247**

Introduction—Aurals and orals in foreign-language teaching—Oral examinations in dentistry and medicine—Simulations—Laboratory practicals—Projects and project work—Projects in subjects other than technology and science—Profiles—Self- and peer assessment

## **12 Objective Tests 285**

Introduction—The language of objective testing—Machine marking and computation—Analysing objective tests—Reliability of objective tests—Test length and test structure—Unusual responses: guessing—Strategies in solving objective items—Intelligent guessing—The question of triviality—Lottery techniques—Item and test construction—Test-wiseness and test preparation—Objective tests in higher education

## **13 Toward Multiple-Strategy Assessment 317**

Retrospect and prospect—From comprehensive to multiple-objective examinations—Case studies—Conclusions

## **Glossary 347**

## **References 361**

## **Author Index 395**

## **Subject Index 407**

## CHAPTER 1

# *Assessment in Higher Education: An Introduction*

## ATTENDING TO ASSESSMENT IN HIGHER EDUCATION: AN HISTORICAL NOTE

When I wrote the first edition of this book in 1977 the title of the opening chapter was inspired by a weekly satire on British television called *Not so much a Programme, More a Way of Life*, for that is what I believed examinations and testing in higher education to be. There was very little interest in Britain in examinations, yet, because they grade individuals, they were and are probably the most important activity in education. Thus in order to create interest in the problem of assessment I began with the following statement, which was typical of the situation at that time:

'Senate approved the syllabus for our new degree course at its last meeting. X will act as Director of Studies and co-ordinate the timetable for the new degree (Head of Department, April 1986).

Please let me have your questions for the Part 1 examination (of the new degree course) by the end of next week at the latest (Director of Studies, February 1987).

Hell, I'd forgotten all about the exam. Let's see what I set in this course in the other degree programme last year ... (Teacher, February 1987).

And the Director of Studies will probably get his questions. The students will be sensible enough to look up the questions set by the lecturer in previous examinations in similar courses, and with a bit more luck, few will fail. Examinations are the great afterthought of the educational process. Most new courses are set up without one thought being given to the methods of examining. The examination is remembered as a necessary evil when the Course Director asks for the annual set of questions and subsequently gets the teacher to vet the printer's proofs of the examination paper. Some teachers use examinations as their objectives (Wiseman, 1961), and some go so far as to give the students the answers in thinly disguised formats. It might be argued in the absence of an understanding of their purposes that this is a step forward, since there is at least a semblance of a relationship between assessment and teaching. Just occasionally someone breaks away from set patterns and involves him- or herself in the additional hard graft which worthwhile innovations in assessment seem to demand.'

Had I substituted 'testing' for 'examinations' I believe that the same or similar paragraph would have been true of the United States. Instead of 'essay' examinations at the end of the year the example would have been informed by the stereotype of multiple-choice tests and quizzes. These were (and generally are still) not common in Britain. There is a great hostility to objective testing, despite its widespread and successful use in medicine, science and technology. It is for this reason that Chapter 12 on objective tests is relatively long in that the arguments for objective tests are set out in detail and recent developments in their analysis described. They are likely to be increasingly used as computer-assisted instruction develops.

At the time that I wrote the first edition the use of the term 'assessment' was relatively novel in Britain. I had wanted to use the phrase 'control of learning' in the title of the book for the reason that I held (and continue to hold) that, if we are to improve learning, we will have to improve the methods of testing and learning we use. They will have to become intimately related. Unfortunately, a book of that title was published just after I had completed the manuscript. For this reason, I used the term 'assessment', which I did not define except by implication.

Until the 1960s the term 'assessment', when used in Britain, applied to the assessment of school pupils for special units which catered for children with special social and psychological needs. Its usage in this way derived from the pioneer of intelligence testing, Alfred Binet, who, in his work in Paris, was concerned with the assessment of mentally retarded children. However, during the 1950s various universities and colleges in higher education (or, more correctly, various individuals in higher education) reacted against the single terminal examination at the end of the year. They felt that it was an inadequate measure of a student's performance, which they believed should be assessed on several occasions during the student's academic career. They called assessment of this type 'continuous'. It was not perceived to be formal, formative assessment (see Chapter 4). Later, in 1969, a major conference sponsored jointly by the Association of University Teachers and the Committee of Vice-Chancellors and Principals of the universities was called 'The Assessment of Academic Performance' (CVCP, 1969). Given this interest in assessment in Britain, I felt that the title was self-explanatory, although some critics did not.

If it was self-explanatory, then it certainly is not now. A study of American literature shows that assessment has a variety of meanings, so it is necessary to examine these in order to resight the parameters of this study. Unfortunately, that is not the only problem of terminology, for, what do we mean by higher education?

When I first chose the title of this book it related in Britain to universities and a limited number of institutions in the public sector called polytechnics, which were able to offer degrees recognized by the State. This is a very limited definition when compared with the definition given to higher education in the United States. These differences often give rise to misleading comparisons between the American and English systems of education.

Its use in that way in Britain arose from the fact that, until 1987, the sector which coped with technician, craftsman and operative qualifications which was financed by the State through Local Education Authorities (LEAs) was called the 'Further Education' sector. Within that sector courses were defined as non-advanced and advanced (DES, 1987b). The advanced courses embraced higher-level technician studies as well as degree-level programmes validated by an officially approved validating authority called the Council for National Academic Awards (CNAA).

Thus when the Society for Research into Higher Education was founded in 1965 it embraced only research done in universities and those parts of the state (public) sector which undertook degree programmes. The activities of the Society reflected the interests of its members, a small body of researchers and university administrators who had been encouraged by the research undertaken for the Robbins Committee on Higher Education (Robbins, 1963) as well as the recommendations of the Hale (1964) and Brynmor-Jones (1964) committees on university teaching methods.

At the same time, appointments were made to posts in higher education *per se*. Most of them were short-period fellowships which undertook research in the areas of residence, methods of teaching, and university examinations. Cox (1967) (whose paper has been much quoted) and this writer were the two appointees in the area of examinations. Nationally, there was very much more interest in teaching methods than examinations, and it is significant that those who made the appointments thought that research on examinations should be conducted independently of the curriculum and teaching. It is a view that continues to be widely held. The picture with which this chapter opened was intended to illustrate this divorce. A primary purpose of this book is therefore to demonstrate that the design of assessment procedures cannot be carried out independently of either the design of instruction or that of the curriculum.

Although there was little interest in research on examinations in comparison with teaching methods and policy, the Society was involved in the publication of two issues of *Universities Quarterly*, which concentrated on topics related to assessment. In the first (published in 1967), in addition to the paper by Cox, another article questioned the twenty assumptions about examinations listed in Table 1 (Oppenheim *et al.*, 1967). They wrote:

'Empirical testing is possible for most of these assumptions, but it will take a long time if it is done at all. In the meantime, the rationality of the examination system could surely be enhanced if universities were to decide which of these assumptions, here made explicit, they wish to maintain and which to abandon'.

Twenty years later despite many innovations, the position was much the same, with the exception that universities worldwide were being forced to explicitly state their views on assessment. Similarly, the belief expressed by Ager and Weltman (1967) in the same issue 'that a variety of techniques should be used in the university examinations, such techniques being chosen according to the functions that the examination performs' is now generally accepted. Whereas in

the past the concern was with the reliability of examinations, interest has now begun to focus on validity and the design of valid measures, especially in the area of professional competency. The meaning of the terms 'validity' and 'reliability' was the subject of scrutiny (Hammersley, 1987). Assessment centre techniques, developed for the selection of managers in industry and commerce, illustrates the new approach to validity in the design of assessment instruments, as do the competency-based approaches to learning used in medicine (Freeman and Byrne, 1976).

Table 1. Twenty assumptions underlying the use of university examinations listed and discussed by Oppenheim *et al.* (1967). (Reproduced by kind permission of the Editor of *Universities Quarterly*)

- 
1. The assumption that university examinations can include some so-called imponderables such as 'quality of mind', 'independent critical thinking', 'breadth', etc. in their assessment.
  2. The assumption that 'quality' of academic performance is rateable on a single continuum from first-class honours to failure.
  3. Whereas many courses include a good deal of practical work, and a few approach some type of apprenticeship training scheme or sandwich course, we usually pay less attention to those aspects of the course when examining.
  4. To some extent it is assumed that examination performance is a mock-real life performance.
  5. The assumption that each examinee should have individual responsibility for his own performance; we do not expect collaboration or teamwork, no matter how common this may be in real life performance.
  6. The assumption that a student who fails has only himself to blame, for not working hard enough, or for being stupid, or in some other way.
  7. The assumption that the proper place for examinations is at the end of certain courses—not later or sooner.
  8. The assumption that university teachers should also be university examiners and university selectors.
  9. The assumptions about the impartiality of examination.
  10. The assumption that the university should have the sole authority to examine at this level.
  11. The assumption that the use of external examiners prevents bias.
  12. The assumption that forced regurgitation of knowledge under stress is predictive of future performance.
  13. Assumptions concerning mental growth and development and the acquisition of an 'educated mind' influence the type and timing of exams.
  14. The assumption that pressure is required.
  15. The assumption that anxiety is necessary.
  16. The assumption that examination results should be distributed in a certain way.
  17. We are forced to make, and then retract, all kinds of assumptions about the comparability of degrees from university to university and from country to country.
  18. The assumption of the need for uniformity in undergraduate exams: all students in a given year group must pass the same examination paper, and we do not allow examinations to be tailored to individual needs.
  19. The assumption that 'learning is to be valued for its own sake' and not merely as a preparation for career and financial gain.
  20. The assumption that the outside world wants the results of university examinations or takes much notice of them.
-

A second symposium published in 1971 dealt with the problem of 'wastage', as it is known in Britain or 'drop-out' in the United States. Apart from a review of research on wastage (Heywood, 1971a), several of the papers were much concerned with the pool of ability in schools which, it was claimed, was not being utilized (for example, E. H. Cox).

These papers showed that wastage could not be understood simply as a function of ability. There were many psychological and social factors which could contribute to performance. For example, some students were found to be temperamentally unsuited to the terminal written examinations which characterized the British system at that time. Others might, for a variety of reasons have chosen the wrong course and given the virtual impossibility of being able to change courses the student was left with no alternative but to leave higher education.

The impact of college life on performance is, as the causal models in Chapter 2 show, complex. Apart from the internal structure of the institution the influences of the sub-system of higher education in which it operates and the system of education more generally are profound. The understanding of and the meanings given to performance are functions of the sub-system observed and depend in no small way on the policies made in respect of participation in the system.

However, having observed the failure of prescriptive approaches to policy problem-solving which treated the sub-educational systems separately, Martin Trow (1974) and this writer in a later report to the Council of Europe advocated open-system studies of the educational system as a whole (see note 2 of Chapter 2).

Policy making should be influenced by a thorough understanding of the social mechanisms at work. For example a highly selective system of higher education will cause the curriculum in secondary (high) schools to be geared toward the entry requirements of higher education. In its turn the system of higher education will both influence and be influenced by the educational requirements of jobs and more particularly the professions. A discernible 'curriculum drift' in which low status institutions emulated high status institutions (e.g. Ivy League and Oxbridge) was noted in both the UK and the US in the sixties and seventies. One outcome in selective systems of higher education is pressure on students in their second-level studies which may cause some of them to experience difficulties with learning. Thus student performance in higher education cannot be isolated from a student's past experience of education and particularly as it caused that student to prepare for his or her future (Figure 1). For this reason, the mechanisms of selection into higher education and their dependence on second-level education are of considerable importance not only to policy makers but to the institution, for they inform teachers' views of what should be studied in higher education and how it should be taught. This applies equally to choice of institution and subject in higher education (Kealy and Rockel, 1987). By implication, as will be demonstrated in the paragraphs on selection which follow, there can be no absolute definition of higher education related to content, since what can be taught not only relates to



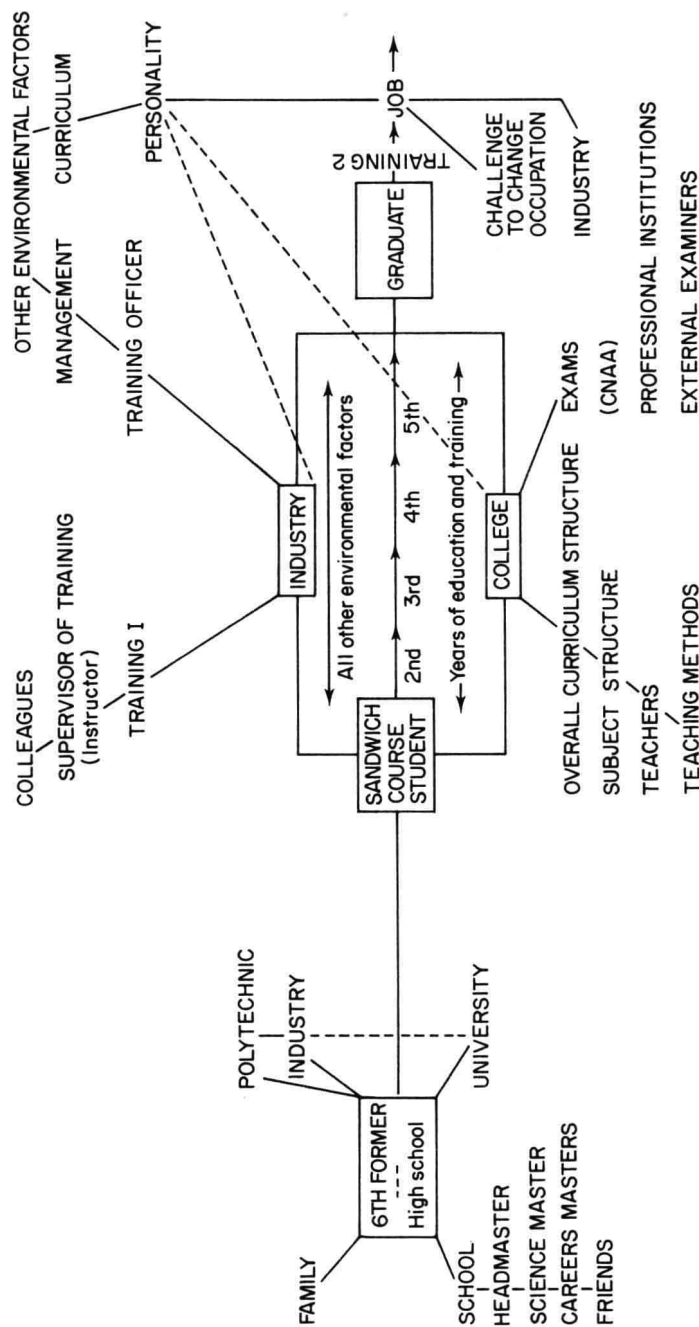


Figure 1. A descriptive model of a subsystem of higher education which shows the forces acting on a sandwich (co-operative) course student in a College of Advanced Technology or polytechnic in Britain. (See Chapter 2 for a discussion of this figure.) (From Heywood, J. (1974c). *New Patterns of Courses and New Degree Courses*, Strasbourg: Council of Europe)