

A HANDBOOK FOR MEDICAL TEACHERS

Third Edition

DAVID NEWBLE, BSc(Hons), MBChB,
MD, FRACP, DipEd

Associate Professor in Medicine
The University of Adelaide
South Australia

ROBERT CANNON, MA(Hons),
MEdAdmin, DipTertEd

Associate Professor and Director
Advisory Centre for University Education
The University of Adelaide
South Australia

Illustrations by **Zig Kapelis**, MArch, MURP,
DipTCP

Formerly Senior Lecturer in Architecture
The University of Adelaide
South Australia



KLUWER ACADEMIC PUBLISHERS
DORDRECHT / BOSTON / LONDON

Distributors

for the United States and Canada: Kluwer Academic Publishers, PO Box 358, Accord Station, Hingham, MA 02018-0358, USA

for all other countries: Kluwer Academic Publishers Group, Distribution Center, PO Box 322, 3300 AH Dordrecht, The Netherlands

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

ISBN 0-7923-8850-X

Library of Congress Cataloging in Publication Data

Newble, David

A Handbook for medical teachers / David Newble, Robert Cannon : illustrations by Zig Kapelis. — 3rd ed.

p. cm.

Includes bibliographical references and index.

ISBN 0-7923-8850-X (case bound)

1. Clinical medicine—Study and teaching—Handbooks, manuals, etc. 2. Medicine—Study and teaching—Handbooks, manuals, etc. I. Cannon, Robert (Robert Anthony) II. Title.

[DNLM: 1. Education, Medical—handbooks. 2. Teaching—methods—handbooks. W 39 M534h 1994]

R834.N49 1994

610'.7'1—dc20

DNLM/DLC

for Library of Congress

94-22258

CIP

Copyright

Third edition © 1994 David Newble and Robert Cannon

Reprinted 2000

First edition published 1983 as *A Handbook for Clinical Teachers*

Reprinted 1984, 1985

Second edition published 1987

Reprinted 1990

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission from the publishers, Kluwer Academic Publishers BV, PO Box 17, 3300 AA Dordrecht, The Netherlands.

Published in the United Kingdom by Kluwer Academic Publishers, PO Box 55, Lancaster, UK.

Kluwer Academic Publishers BV incorporates the publishing programmes of D. Reidel, Martinus Nijhoff, Dr W. Junk and MTP Press.

Printed and bound in Great Britain by Butler and Tanner, Frome and London.

INTRODUCTION

Being involved in student assessment is a responsibility for all of us. In fact, it is the responsibility of all of us who are involved in teaching. It is not just the responsibility of the teacher, but it is also the responsibility of the administrator, the parent, and the student. The purpose of this book is to provide you with the information you need to understand the importance of assessment and to help you develop a plan for assessing your students. This book is designed to be a practical guide for you, whether you are a teacher, administrator, or parent. It will provide you with the information you need to understand the importance of assessment and to help you develop a plan for assessing your students. This book is designed to be a practical guide for you, whether you are a teacher, administrator, or parent. It will provide you with the information you need to understand the importance of assessment and to help you develop a plan for assessing your students.



THE PURPOSE OF THE ASSESSMENT

The purpose of the assessment is to provide you with the information you need to understand the importance of assessment and to help you develop a plan for assessing your students. This book is designed to be a practical guide for you, whether you are a teacher, administrator, or parent. It will provide you with the information you need to understand the importance of assessment and to help you develop a plan for assessing your students. This book is designed to be a practical guide for you, whether you are a teacher, administrator, or parent. It will provide you with the information you need to understand the importance of assessment and to help you develop a plan for assessing your students.

Understanding the purpose of the assessment is the first step in developing a plan for assessing your students. This book is designed to be a practical guide for you, whether you are a teacher, administrator, or parent. It will provide you with the information you need to understand the importance of assessment and to help you develop a plan for assessing your students.

- Understanding the purpose of the assessment
- Developing a plan for assessing your students

7: ASSESSING THE STUDENTS

- Evaluating the effectiveness of the assessment
- Developing a plan for assessing your students
- Understanding the purpose of the assessment

Thought it may be possible for one assessment method to achieve more than one of these purposes. If the information you need is used for appropriate purposes and the assessment is designed to provide valid and reliable data, then it is possible for one assessment method to achieve more than one of these purposes.



INTRODUCTION

Being involved in student assessment is perhaps the most critical of all tasks facing the teacher. Generally, teachers take such involvement quite seriously but, sadly, the quality of many assessment and examination procedures leave much to be desired. The aim of this chapter, therefore, will be to help you to ensure that the assessments with which you are involved will measure what they are supposed to measure in as fair and as accurate a way as possible. We will provide some background information about the purposes of assessment and the basic principles of educational measurement. We will then detail the forms of assessment with which you should be familiar in order that you can select the best method to use.

THE PURPOSE OF THE ASSESSMENT

When faced with developing an assessment you must be quite clear about its purpose. This may appear to be stating the obvious but try asking your colleagues what they think is the purpose of the assessment with which you are concerned. We are certain that there will be a considerable diversity of opinion. Some may see it as testing the students' mastery of the course content, others may see it as a way of ranking the students, and yet others as a way of encouraging students to study the content of their course.

Mehrens and Lehmann identify several purposes of assessment which may be paraphrased as follows:

- Judging mastery of essential skills and knowledge.
- Measuring improvements over time.
- Ranking students.
- Diagnosing student difficulties.
- Evaluating the teaching methods.
- Evaluating the effectiveness of the course.
- Motivating students to study.

Though it may be possible for one assessment method to achieve more than one of these purposes, all too often assessments are used for inappropriate purposes and consequently fail to provide valid and reliable data.

It must never be forgotten how powerfully an assessment affects students, particularly if it is one on which their future may depend. This influence may be a positive one or a negative one and even a harmful one. For many students, passing the examination at the end of the course is their primary motivation. Should this examination not be valid, and thus not truly reflect the content and objectives of the course, then the potential for serious distortions in learning and for making errors of judgement about students is evident. An example from our own experience may illustrate this point. A revision of the final year curriculum inadvertently led to the multiple-choice test component of the end of year assessment having considerably more weight than the clinical component. Students were observed to be spending excessive amounts of time studying the theoretical aspects of the course in preference to practising their clinical skills, the latter being the main aim of the curriculum revision. A subsequent modification of the assessment scheme, giving equal weighting to an assessment of clinical competence, has corrected this unsatisfactory state of affairs.

It is our view that assessments on which decisions about the students' future are to be made (summative assessments) should be kept quite separate from assessments which are for the benefit of the student in terms of guiding their further study (formative assessments).

Summative assessment

In dealing with summative assessment, every effort must be made to ensure that all tests are good tests so that any decisions made are fair and based on the appropriate criteria. Students should be fully informed of these criteria, on the test methods to be employed and on the weightings given to each component. Such information should preferably be given to students when a course begins. This is important because it is surprising how often information obtained from other sources, such as past students or even from the department itself, can be inaccurate, misleading or misinterpreted by the students. The best way of avoiding this is to provide printed details of the course plan, including the assessment scheme. Examples of past papers can be provided and we have found an open forum on the assessment scheme early in the course to be both popular and valuable.



Formative assessment

Formative assessments can be organized more informally. Such assessments must be free of threat as the aim is to get the students to reveal their strengths and weaknesses rather than to disguise them. Opportunities to obtain feedback on knowledge or performance are always appreciated by students and lead to positive feelings about the department and the staff concerned.

WHAT YOU SHOULD KNOW ABOUT EDUCATIONAL MEASUREMENT

Whatever the purpose of the assessment, the method used should satisfy the following three requirements:

- 1 Validity:** Does it measure what it is supposed to measure?
- 2 Reliability:** Does it produce consistent results?
- 3 Practicability:** Is it practical in terms of time and resources?

For example, a test of blood sugar would be invalid if it actually measured blood alcohol. It would be unreliable if the sample of blood gave different results on repeated analyses. It would be impracticable if it took two technicians three hours to perform at a cost of \$500 per test. This example may be unrealistic but unfortunately, analogous education examples of invalid and unreliable tests can be found in most, if not all, medical schools. For instance, the results of multiple-choice examinations are frequently used as a major component of examinations which are used to certify clinical competence. Though such tests may be high on reliability they are predictably low on validity. Another example is the traditional clinical viva which continues to be widely used despite the evidence showing that the inconsistent performance of the examiners and the small sample of observed student behaviour makes such assessments very unreliable.

Our intention in raising these matters is to encourage you to apply the same critical interest in the quality of educational tests as you undoubtedly apply to those tests you use in your research or in the investigation of your patients. This section will provide you with some basic information about aspects of educational measurement with which we think you should be familiar.



Validity

Content validity is the first priority of any assessment. It is a measure of the degree to which it contains a representative sample of the knowledge and skills it was meant to cover. A numerical value cannot be placed against it and it must be subjectively judged according to the objectives of the assessment. Therefore in approaching any assessment the first question you must ask is: **what are the objectives?**



Unfortunately, in many medical schools such objectives are not available, either for the institution or for individual courses. Should you be in this situation, with no written objectives for the assessment you have to design, then you have no alternative but to develop your own. This is not such a difficult task as you might imagine because, as far as the assessment is concerned, the objectives are embodied in the course content. A look at the teaching programme, lecture and tutorial topics, and discussions with teaching staff should allow you to identify and categorize the key features of the course. What you are, in fact, attempting to do is to construct a course plan in reverse and you may find it helpful at this point to consult the chapter on course design where this process is discussed in greater detail.

The objectives of the course, however categorized, are the framework against which you can evaluate the content validity. For the content validity to be high, the assessment must sample the students abilities **on each objective**. As these objectives will usually cover a wide range of knowledge, skills and attitudes, it will immediately become apparent that no single test method is likely to provide a valid assessment. (For instance, a multiple-choice test cannot provide valid information about manipulative skills.) This sounds simpler in theory than it is in practice. Some objectives, particularly those in the attitudinal area, will be hard to assess in an examination setting. In such cases, alternative forms of assessment should be sought. For example, you might have to agree that the attitudinal objectives of the course should come from supervisors' ratings despite the problems you anticipate with their reliability.

There are two other forms of validity known as **construct validity** and **criterion-related validity**. Generally speaking you will not be in a position to evaluate these aspects of

validity so we will not discuss them further.

Reliability

The reliability of any assessment is a measure of the consistency and precision with which it tests what it is supposed to test. Though its importance is initially less vital than validity, you should remember that an unreliable assessment cannot be valid. The degree of reliability varies both with the test format itself and with the quality of its application.

Theoretically, a reliable test should produce the same result if administered to the same student on two separate occasions. The statistic describing the degree to which this happens is called the coefficient of stability. In most situations, practical considerations make it impossible to provide such information. This difficulty is normally overcome by providing a measure of **internal consistency**. The basic principle of this statistical technique is to split the test into two parts and assume they are equivalent. A test with a high degree of internal consistency shows a high correlation between the student's performance on each half of the test. The most commonly used statistics of this type are the Kuder-Richardson formulae. You will frequently come across them in computer scored multiple-choice tests.

The other key component in determining the reliability of a test is the **consistency of the marking**. This factor is responsible for the unacceptable levels of reliability in most forms of direct assessment, in clinical examinations and in written tests of the essay type. However, methods are available to help you minimize this problem and will be discussed later in this chapter.

Norm-referenced versus criterion-referenced assessment

You are likely to be familiar with norm-referenced assessment as this reflects the traditional approach to testing. Any assessment which uses the results of all the students to determine the standard is of this type. In such tests the pass level is often determined arbitrarily, by pre-determining the proportion of students given each grade



or statistically, by using the standard deviation.

In the professional field we should be more concerned that the students achieve some minimal standard of competence. If this is the main purpose of the assessment then the criterion-referenced approach is the more appropriate. Such an approach necessitates the determination of the standard before the examination is conducted rather than waiting to see the overall results before doing so. Though this is extremely difficult to do, and even then often has to be modified for practical reasons (e.g. too many students would fail), we have found that attempting to do so is a powerful way of improving the validity of the assessment. Everyone concerned is forced to consider each item in the examination and ask themselves if it is relevant and set at the appropriate level of difficulty. Our own experiences with such an examination used to test clinical competence in the final year have been very rewarding.

ASSESSMENT METHODS

In planning your assessment, it is clearly necessary to be aware of the variety of methods available to you. It is impossible to be comprehensive for reasons of space so we will restrict ourselves to those methods most commonly in use in medical schools. We will also include information about some of the more innovative approaches which have been developed recently which may be of interest. We do this deliberately in an attempt to encourage you to experiment. As you look more critically at assessment procedures you will soon face situations where it is obvious that inappropriate methods are being used. This may be due to a combination of tradition, ignorance and prejudice. The first two you may be able to influence by rational argument based on the type of information we provide in this book. The latter is a more difficult problem with which to deal.

PROFESSORS OF MEDICINE AND SURGERY
BY A COMMITTEE OF THE HEALTH CENTRE
SELECTION OF SERVICES AND PROVIDE ILLUSTRATIONS FROM
CENTRES OPERATING IN AT LEAST TWO DIFFERENT REGIONS.