

Handbook on  
Testing and  
Evaluation  
in Business Education

# Handbook on Testing and Evaluation in Business Education

**Margaret J. Macdougall**  
Professor  
Division of Vocational Education  
Faculty of Education  
University of New Brunswick  
Fredericton, New Brunswick

**Gerald L. Roussie**  
Associate Professor  
Division of Vocational Education  
Faculty of Education  
University of New Brunswick  
Fredericton, New Brunswick

Editorial Development: Marion Elliott



**Copp Clark Pitman**  
A division of Copp Clark Limited  
Toronto

©Copp Clark Pitman 1983

All rights reserved. No part of the material covered by this copyright may be reproduced in any form or by any means (whether electronic, mechanical or photographic) for storage in retrieval systems, tapes, discs, or making multiple copies without the written permission of the publishers.

ISBN 0-7730-4320-9

Editing/Penny McIraith  
Design/Riitta Malm  
Cover/Patti Brown  
Typesetting/Kerr Graphics Ltd.  
Printing/The Hunter Rose Company Ltd.  
Binding/The Hunter Rose Company Ltd.

## Acknowledgements

Copp Clark Pitman, for a letter from *Canadian College Typewriting* by Sandra Ubelacker et al, ©1979.  
Gage Publishing Limited, for an evaluation sheet from *Office Procedures 2000, Instructor's Guide and Manual* by Douglas G. Westgate, ©1977.  
Geyer-McAllister Publications, Inc., for an excerpt from "Measuring Repetitive Office Work," by Robert L. Caleo in *Office Administration and Automation*, ©1964.  
McGraw Hill Inc., for Letters 6, 18, 19, and 56 from *Dictation for Mailable Transcripts* by Louis A. Leslie and Charles E. Zoubek, ©1950, and for two letters from "Take 78," from *Speed Dictation with Previews in Gregg Shorthand* by Charles E. Zoubek, ©1963.  
Prentice Hall Inc., for the classification of learning and test types from *Essentials of Educational Measurement 2nd ed.*, p. 65, by Robert I. Ebel, ©1972.

## Canadian Cataloguing in Publication Data

Macdougall, Margaret J.

Handbook on testing and evaluation in business education

Bibliography: p.  
Includes index.

ISBN 0-7730-4320-9

1. Business education. 2. Business education - Examinations. I. Roussie, Gerald L. II. Title.

HF1121.M32

650'.076

C83-098598-0

Copp Clark Pitman  
495 Wellington Street West  
Toronto, Ontario  
M5V 1E9

Printed and bound in Canada

---

## Preface

Three of the most common issues facing business teachers are: (1) a recognition of the various components of office tasks and the standards required for success on the job, (2) a reconciliation of business classroom standards with realistic job standards, and (3) the selection of the best instrument for evaluating specific business subjects at various stages of student development. These issues are addressed in *Handbook on Testing and Evaluation in Business Education*.

The Handbook begins with the authors' philosophy of vocational business education. This is followed by a discussion of the role of evaluation in ensuring that the results of ongoing job analyses are used to maintain curriculum relevancy and appropriate classroom standards. The foundation for a sound testing program is outlined in the first two chapters, while the third chapter discusses types and uses of tests.

Two major types of tests, *written* and *performance*, form the basis of evaluation in a comprehensive business program. Using such subjects as communications, business law, business organization and management, economics, general business, and others as examples, written tests are discussed and exemplified. Performance tests such as those associated with accounting, shorthand, typing, and office procedures are given detailed coverage in separate chapters.

The authors realize that the process by which teachers evaluate classroom learning is subject to change, but emphasize that the principles of evaluation remain constant. Consequently, *Handbook on Testing and Evaluation in Business Education* is adaptable to technological changes.

The material in this handbook is the result of many years of experience in the preparation of students for the office job market and/or the business classroom in addition to the feedback from employers in both business and educational fields. As such, it should be useful to the novice and experienced teacher in the preparation of vocationally competent workers for the modern office.

To all students, colleagues, and employers who directly or indirectly have made a contribution, a sincere thank you. Sincere appreciation is extended to Marion Elliott and Penny McIlraith for their excellent editorial guidance.

M.J.M.  
G.L.R.

---

# Contents

	Page
PREFACE	iv
CHAPTER 1 Philosophy of Testing and Evaluation	1
CHAPTER 2 Elements of a Valid Test	9
CHAPTER 3 Types and Uses of Tests	13
CHAPTER 4 Testing in Basic Business Subjects	36
CHAPTER 5 Testing in Accounting	46
CHAPTER 6 Testing in Typewriting	67
CHAPTER 7 Testing in Shorthand	123
CHAPTER 8 Testing in Office Procedures	186
CHAPTER 9 Testing: A Means, Not an End	198
BIBLIOGRAPHY	201
INDEX	204

# Philosophy of Testing and Evaluation

## Introduction

Educators' philosophies of education will influence their goals and objectives. These, in turn, will influence the materials and procedures used in the classroom. Finally, the materials and procedures used will be reflected in the evaluation procedures. When the underlying philosophy of a program is the preparation of students for jobs in business, it is necessary to determine the needs of the modern office and the worker and to specify terminal performance objectives which are in accordance with job requirements. In other words, when the aim is to prepare business education students who are vocationally competent, there must be a means of determining whether or not the student *is* vocationally competent. The minimum terminal performance objective of the classroom must be at least equal to the minimum initial job requirements or performance standards.

The primary reason business education came into the secondary and post-secondary classrooms was to provide vocational education—to prepare students for essential jobs in the work force. Despite words to the contrary, job preparation remains the primary reason for the continued support of business education. Students electing to follow a business curriculum should, at the end of the training, be prepared to enter the business labor force as competent workers in the particular area for which they were trained. The level of competency will be influenced by the level of training. In this text, the concern is with students entering the clerical work force from secondary school systems or from the post-secondary or community college institutions.

Determining vocational competency standards may be a difficult, frustrating exercise. Immediately the subject is raised, numerous problems surface. Standards may vary according to the availability of workers and jobs, or they may vary from one geographical region to another. Information based on appropriate time studies or work measurement may not be readily accessible to the classroom teacher. Although such conditions may exist, they do not constitute a valid reason for neglect.

The professional teacher who is genuinely interested in current and future students will not ignore the responsibility of coming to grips with the problem and deciding, for a particular situation, what the minimum accepta-

ble vocational competency standards will be. Unless vocational standards are set by the classroom teacher and the school system, the employer has no assurance that students will possess saleable vocational skills at the conclusion of their business program; students may possess a school certificate which has low marketable power in the business market.

This text is concerned with specific suggestions for evaluating performance in business education subjects. It reflects the authors' philosophy concerning the goals and objectives of vocational business education programs. Consequently, the suggestions are an integral part of and reflect their individual teaching procedures. Before examining these suggestions, the reader needs to be cognizant of the underlying bias and philosophy of the writers concerning the topic of evaluation and its relationship to the responsibility of the business teacher who seeks to strike a workable balance between the needs of the work force and those of the student.

## Role of Evaluation

Evaluation is an essential part of the ongoing learning-teaching process. In the learning-teaching process, there are definite, distinct, and often overlapping stages or steps. Through the years, these have been described in diverse ways. Individuals have used varying descriptive words, but the essential ideas are constant.

Tonne (p. 29) referred to Morrison's steps in the Mastery Formula as:

1. Pretest.
2. Teach.
3. Test the results.
4. Adapt the procedures.
5. Reteach.
6. Test again, and
7. Continue to the point of actual learning.

Evaluation (test the result, test again) was considered to be part of the process. Without evaluation there was no basis on which to make decisions for subsequent procedures.

## Steps in Evaluating

Clark and Starr (p. 4) listed and described four steps in the teaching-learning process:

1. Diagnosing the learning situation.
2. Preparing the setting for the learning.
3. Guiding the learning activities, and
4. Evaluating the pupil's learning.

With one major exception, these, in essence, were the same steps as those described by Morrison. The exception was that nothing seemed to be done with the test results. The testing could almost have been an end in itself rather than a means to an end — mastery of what was to be learned.

In the second edition (p. 11) of this text, this oversight was corrected by the addition of a fifth step — follow-up.

Current emphasis in learning focuses on such terms as job analysis, competency-based instruction, and criterion-referenced instruction. Consideration of the work of such people as Bloom, Block, and Carroll indicates that in teaching or learning, the first step is ascertaining present position, i.e., pretesting. Next comes deciding the destination, i.e., goal setting. Finally, a decision as to the means of knowing whether or not the destination has been reached must be made, i.e., evaluation.

Those steps listed by Morrison and expanded by Clark and Starr have been rewritten (Figure 1) to include job analysis, recognition of office requirements, and provision for ongoing revision of curriculum and standards.

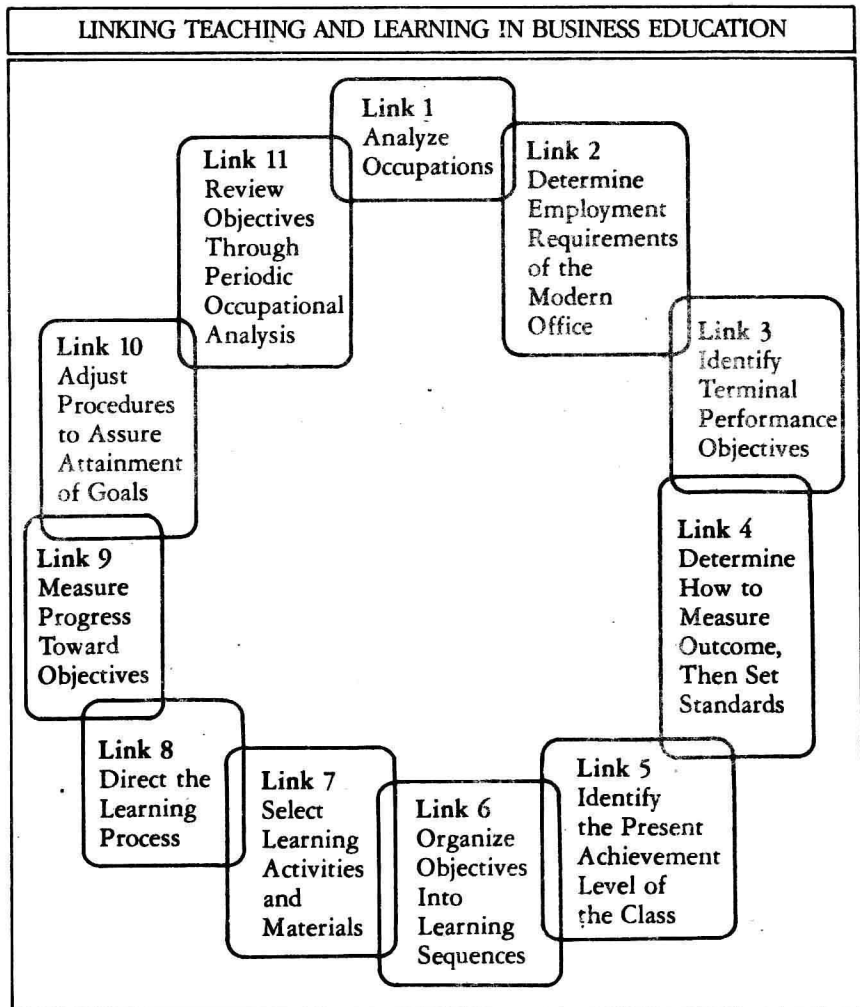


Fig. 1



Figure 1 illustrates the writers' concept of the teaching-learning strategy in preparing students for office employment. The major link in the chain is Link 2, "Employment Requirements of the Modern Office." To meet the employment needs of this modern office, each job must be analyzed (Link 1), terminal performance objectives must be identified (Link 3), a method of measuring outcomes must be determined and specific standards set (Link 4), and mastery tests must be prepared (Link 9).

It is then necessary to determine where the student stands with respect to objectives and related materials (Link 5). As the student progresses, it may be necessary to adjust the instructional program to ensure that performance objectives will be achieved (Link 1C). Finally, to ensure that the classroom does not become irrelevant to the needs of the modern office, there must be a periodic review of objectives through an ongoing occupational analysis (Link 11). If the subsequent occupational analyses show a change from the previous ones, terminal performance objectives will have to be revised accordingly.

## Evaluation Versus Measurement

In this handbook, these words, although closely related, are not used interchangeably. Measurement is limited to *quantitative* descriptions of student behavior. It does not include qualitative descriptions, nor does it imply judgments concerning the worth or value of the behavior measured. It answers the quantitative questions, "How much?" or "How many?"

Evaluation, in addition to being concerned with these same questions, is also concerned with answers to the question, "What kind?" and "How well?" It includes *qualitative* as well as *quantitative* descriptions of student behavior plus value judgments concerning the ability of the student to display this behavior. The medical doctor counts the pulse, reads the temperature, notes blood count, and weighs and measures according to all available instruments. By pondering the results of all measurements and other observations, a value judgment concerning the health of the patient is reached. In like manner, the professional teacher measures, counts, and notes while observing and guiding the students' activities and behavior to determine the present status of the student.

*Measurement* asks "How many apples are in the box?" Consequently, the apples are counted or weighed and a numeric result is obtained. This is where measurement stops. Evaluation begins when we ask, "What kind of apples are they? Are they red apples? If so, how red are they? Are they juicy? If so, how juicy? What is the taste? How do these apples compare with those in the other box?" Certainly apple producers may have developed scales which will give numeric-type answers to many of these questions. But eventually all these answers must be viewed collectively and a value judgment made as to the quality of a specific apple or number of apples.

## Importance of Evaluation

The importance of evaluation lies in the use made of the results — the use made by the student, the teacher, administration, and the employer. In Figure 1, the teaching-learning chain, it is specifically Link 9 "Measure Progress Toward Objectives" where one pauses to take stock of progress toward goals. Evaluation is also part of the ongoing learning activities giving student and teacher immediate knowledge of results, the means of determining whether or not the student is on the right track. As such, it ought, for the student, to be a strong motivating device — a spur to keep the student actively involved. To the teacher it is the means of obtaining information on which to base decisions. To achieve the goals associated with measurement and evaluation, the teacher will use formal and informal testing procedures.

## Why Test?

There are two major reasons for testing:

1. To determine how the students' performance compares with that of similar groups of students (normally referred to as norm-referenced testing).
2. To determine how much progress the student has made toward the attainment of specific terminal performance objectives (criterion-referenced testing).

It is with the ninth link of the teaching-learning chain, "Measure Progress Towards Objective," that this text is most concerned.

The measurement chain shows that there must be specific terminal performance objectives for any course. Eventually, it is necessary to determine to what degree students have attained the skills or knowledge outlined in the performance objectives. Thus, one answer to the question, "Why test?" is quite evident: there must be some way of indicating student growth.

The ninth link in Figure 1, "Measure Progress Toward Objectives," indicates that the classroom teacher must measure the attainment of performance objectives. The outcome of this link will be the adjustment of the instructional programs to ensure the achievement of performance objectives. If the teacher avoids this step, neither the teacher nor the student will have a measuring tool by which to determine whether to continue existing practices or to change approaches to instruction. Thus, one of the functions of a test is to provide both parties with sufficient knowledge on which to base a decision.

## Tests to Determine Student Readiness

The fifth link in the teaching chain "Identify the Present Achievement Level of the Class" requires an analysis of a student's readiness and preparation to enter a specific area of learning. When the teacher analyzes the objectives in relation to a student's preparation, a decision can be made as to what level

of preparation that student has already reached. An illustration of this is the necessity for a teacher to determine the performance level of a transfer student. It is in this testing situation that students can be classified according to their present level of achievement so that readiness may be determined.

There may also be instances within a particular course in which a teacher learns of a student's ability to perform a certain task because of some previous experience. Suppose a student in a general business class tells the classroom teacher that he or she already knows the unit on banking, and in particular bank deposits and reconciliation statements, because the student has been doing these tasks in his or her parent's service station business. Once the teacher has determined that this student can perform the task to the same degree as outlined in the objectives for that unit, alternative learning situations may be sought for the student. If a student has been able to perform adequately for a teacher prior to instruction in a particular area, then it is the teacher's responsibility to provide enrichment for the student rather than require repetition of a task already mastered.

## Tests to Assist in the Guidance Function

Determining student readiness has a relationship to the guidance function of the classroom teacher. If a student has the aspiration to seek a career in a given field, the teacher can provide that career information in relation to the student's readiness to accept the responsibilities of the new or advanced field of study. One such example would include the student who has aspirations of working toward a CGA or RIA designation because of an interest and/or ability in accounting. On the basis of the teacher's measurement of the student's ability in class, the teacher can assist the student in working toward a career in the accounting area. It may be possible to guide the student into an alternative course of study if the teacher feels that the student's choice is too far out of reach. On the basis of having measured the student's capabilities over a year's work or a term's work, the teacher can identify the student's state of readiness to advance in related course work.

The secretarial student can also be guided into new career paths as a result of those tests that determine readiness. Suppose a student indicates an interest to enter a career in court and conference reporting and expresses this aspiration to the shorthand teacher. On the basis of the student's past performance, the teacher may be able to point out that the student may not possess the potential or preparation for such a career. Consequently, it would be far better for the student to understand the state of readiness and improve skills and knowledge before embarking on a career that would prove to be too advanced for the present state of readiness.

## Tests as Motivators

When a test has been assigned by a classroom teacher, and the student knows the scope of the material to be tested, that test can act as a positive

motivating force. Preparation by the student requires selecting the major concepts from the unit to be tested and combining them into cohesive units. The additional effort expended in preparing for the test acts as a reinforcer, thus giving the student a more thorough understanding of the major concepts required. The motivating factor is carried even further when the test is based on those concepts outlined in the performance objectives for the unit.

An additional example of motivation through testing can be found in those instances in which the classroom teacher provides reinforcement to the student through immediate feedback. When the student has been given the correct response to a test question and when the teacher has given an adequate explanation of that response, the student is given a further opportunity for understanding the concept tested. That additional explanation could result in a student's feeling of accomplishment because the concept at last has been understood.

The tests — and more particularly the results of the test — can prove to be a motivating factor for the student, especially if there have been successes on the test itself. When the test is tied directly to the objectives and when the student is able to prepare for the test on these same bases, then there is the opportunity to do well — indeed a motivating factor in itself. The fear of the unknown (unspecified test objectives) will be eliminated when the student knows the basis for the test content, and the absence of this fear acts as a positive motivator.

## Tests That Assist the Teacher

Tests may have an impact on the teaching function when the teacher uses the test results to separate students for instructional purposes into groups that are relatively homogeneous in abilities. The grouping of the students does not have to be a physical action, but can be a mental process that the teacher uses in preparation for the class. One example of this would involve the shorthand teacher who gives a five-minute dictation to the class. After grading the transcript, the teacher can then group the class into different speed groups so that dictation at varying levels can be given.

## Summary

The aim of this chapter was to stress the importance of testing in the teaching-learning process and to identify specific functions of tests. Graphically illustrated as a chain, the strategy began with an occupational analysis that would reveal the employment requirements of the modern office. Terminal performance objectives then may be set that would be tied directly to the business office. The fourth link in the teaching-learning chain emphasized the importance of deciding the most appropriate way of measuring the outcome of the objectives and setting standards for the course being planned.

Before the teacher begins to teach, it is essential that an appropriate starting point be determined. Once the present achievement level of the

class has been learned, the teacher can begin to organize the objectives into learning sequences, select the proper learning activities and materials, and then direct the learning process. It is the next link that provides the greatest strength in the chain, that of measuring progress toward objectives. Unless the teacher and student have planned for some method of determining how much progress has been made toward reaching the end goals, there is a possibility that both may wander aimlessly. Testing and evaluation must be an integral part of that plan.

Once the testing has been undertaken, the teacher can adjust procedures accordingly and either retrace some steps to better assure attainment of the goals or advance toward newer goals.

Since this is an ever-changing world, and since new procedures and methods are making their presence known through advanced technology, it is extremely important that the occupational analysis be an ongoing process. From the results of a subsequent analysis, up-to-date information about employment requirements of the modern office will lead to a review of terminal performance objectives.

This chapter also outlined the various functions of tests and elaborated on their usefulness in determining student readiness, in assisting in the guidance function, as motivating factors, and in providing assistance to the teacher.

# Elements of a Valid Test

E.L. Thorndike has outlined two types of educational measurement:

1. That which measures how well a student does something as compared to how other students of similar backgrounds do it (*norm-referenced*).
2. That which measures how well and how hard a task a student does as compared to a previously established standard (*criterion-referenced*).

## Criteria for Test Construction

A precise vocabulary pertaining to desirable test characteristics has been developed. Though developed primarily in connection with standardized or norm-referenced testing, such qualities as relevancy, reliability, objectivity, adequacy, and validity are equally applicable to criterion-referenced testing. The maker of classroom tests cannot ignore these qualities and the need to include them in all tests. A brief overview follows.

*Relevancy* refers to the relationship between what is required in the testing process and the purpose of the learning.

*Reliability* refers to the consistency with which similar results are obtained from repeated use of the test or parallel forms of the test.

*Objectivity* in test construction means that the response to a test item or task would be identical from experts in the field.

*Adequacy* is concerned with the length, number, and difficulty of test items.

When there is no doubt as to the interpretation of the test item (objectivity), when that item is in line with the stated purpose of the learning (relevancy), when it is of sufficient length and difficulty (adequacy) to insure consistent responses (reliability), the test will possess the culminating or all-inclusive characteristic of *validity*.

## Evaluating Test Construction

**Illustration 1** During the first unit in journalizing in an introductory accounting class, the basic task to be mastered was stated as the ability to journalize, in a ten-minute period with 90 per cent accuracy, ten representative transactions of a service business.

An accounting test consisting of ten typical transactions with similar wording to that used in the practice exercises was prepared and administered. Approximately 90 per cent of the students scored 90 per cent or higher. Were the test items *relevant*, *reliable*, *objective*, *adequate*, and, consequently, *valid*?

*Critique:*

1. **Relevancy.** A comparison of test items with stated objectives indicates that the test asked students to do what was stated in the objectives.
2. **Reliability.** It was possible to give the same test to similar students; results were approximately the same, or consistent.
3. **Objectivity.** An examination of student errors did not indicate that errors were due to a misunderstanding of the written narrative of transactions.
4. **Adequacy.** The test items appeared to be of equal length and difficulty as those in the practice exercises and could be said to be adequate.

*Conclusion.* The test was in line with the stated objectives; it was of the right length and difficulty; there appeared to be no problem of interpretation; the results obtained were consistent. It can be assumed that the test is valid for this particular objective.

**Illustration 2** During a first unit in journalizing in an introductory accounting class, the basic task to be mastered was stated as the ability to journalize in a ten-minute period with 90 per cent accuracy, ten representative transactions of a service business.

An accounting test consisting of twenty transactions to be journalized in ten minutes was prepared. The transaction descriptions were lengthy, contained many low-frequency words, involved sentence structure, superfluous information, and were open to varying interpretations. The majority of the students scored below 50 per cent. Were these grades due to lack of knowledge of journalizing procedures at the stated standard or were these low grades due to the use of an invalid test? Were the test items *relevant*, *reliable*, *objective*, *adequate*, and, consequently, *valid*?

*Critique:*

1. **Relevancy.** The objective specified transactions relating to a service business while the test question dealt with transactions pertaining to a merchandising business.
2. **Reliability.** This can only be determined by administering the same test on several different occasions to similar groups of students. It is conceivable that the results of subsequent testing may yield similar results. Thus, the test might be considered reliable.
3. **Objectivity.** Answers given for some of the transactions led to the conclusion that the students did not interpret the narrative similarly.
4. **Adequacy.** The narrative for each transaction was long, involved, and ambiguous. These violations had a negative effect on the relevancy of the test question.

*Conclusion.* There is the absence of the essential characteristics of relevancy, adequacy, and objectivity. Consequently, the test is invalid.

## Effects of Educational Goals on Method of Assessment

Within a balanced business curriculum, the goals of specific learning-teaching situations will cover all the levels of learning outlined for the cognitive, affective, and psychomotor domains of learning and will require different teaching-learning strategies. When learning strategies are different, testing strategies may also differ. Ebel's (p. 65) classification in Figure 2 focuses on this.

An examination of Ebel's chart indicates that the written test is a recommended method of assessment in each of the basic areas of learning. If this is so, it is essential for the classroom teacher to be aware of the range of written tests available, of the strengths and weaknesses of each, and of guidelines to be followed in the preparation and use of each. Written tests may be divided into three groups:

1. Short-answer tests including such types as completion, true-false, multiple choice, and matching.
2. Longer-answer supply tests such as listing, describing, or defining.
3. Essay tests such as composition of a letter or the analysis of a case study.

Chapter 3 suggests procedures to follow for construction, administration, and scoring of each of these types of written tests.

<b>Educational Goal</b>	<b>Learning Process</b>	<b>Method of Assessment</b>
<b>Knowledge</b> Concepts Facts Reasons Processes	Observing Reading Reflecting Expressing	Written test Incidental observation
<b>Skill</b> Writing Speaking Typing Drawing Swimming, etc.	Observing Imitation Practice Knowledge acquisition	Performance Written test Incidental observation
<b>Character</b> Personality Habits Attitudes Values Standards	Observing Imitation Conditioned behavior Knowledge acquisition	Incidental observation Written test

Fig. 2 *Methods of assessment according to educational goals and learning process used.*



## Underlying Principles of Test Construction

Before a teacher can proceed to prepare a valid test, certain principles should be considered:

1. **Need for unambiguous goals.** In the learning-teaching situation, an imperative is the careful statement of objectives so that the learner's perception of the task runs parallel to the teacher's perception of the task. The closer the learner's perception of the task is to the teacher's perception of the task, the greater should be the degree of pertinent learning.  
It follows that clearly stated objectives should leave no room for doubt in the learner's mind as to the task to be accomplished. It remains only for the teacher to construct test items in such a way that they clearly reflect the stated objectives, lack ambiguity, and produce a learning product which can be readily and easily compared to the established criterion.
2. **Need for tests to be prepared at the time objectives are prepared.** If this is done, lack of clarity in objectives and/or test items can be identified and corrected.
3. **Need for the teacher to check the completed test against the criteria of relevancy, reliability, objectivity, adequacy, and validity.**
4. **Need for the teacher to give careful consideration to the administration and scoring plans for the test.** Of greatest importance will be the procedures followed to foster development of a healthy student attitude toward the testing process. Good testing procedures ought to produce students who recognize tests for what they should be — learning tools designed to help reach desired objectives.
5. **Need for the test maker to prepare and evaluate this test on the basis of three important attributes** (a) knowledge of subject matter, (b) knowledge of standards necessary for the competency level required, and (c) the appropriate selection of test items.
6. **Need for test items to be concerned with the important concepts studied.** Weight to be proportionate to amount of class time spent developing each particular concept.
7. **Need for the classroom teacher to be familiar with the types of available evaluation instruments; uses and limitations of each type; and guidelines for the construction, administration, and scoring or grading of each type.**

## Summary

This chapter was concerned with those elements necessary in a valid test. To ensure that a test is relevant, it must contain items which are directly related to the objectives of the unit being tested. A test composed of an appropriate number of items of suitable difficulty and length possesses adequacy. Unambiguous statements should eliminate erroneous interpretations and contribute to objectivity. Comparison of results on additional administrations of the test to similar groups will indicate the presence of reliability. Unless a test possesses all of these characteristics, it is not valid and may yield little useful information in the teaching-learning process.