DOCUMENT RESUME

ED 156 556 95 SO 010 776

TITLE Evaluation in Consumers' Education: Teacher Education Module 4.
INSTITUTION Florida State Dept. of Education, Tallahassee.
SPONS AGENCY Bureau of Occupational and Adult Education (DHEW/OE), Washington, D.C. Office of Consumers' Education.
PUB DATE [77]
NOTE 76p.; For related documents, see SO 010 773-775
AVAILABLE FROM Florida Department of Education, c/o ITV and Educational Products Distribution, Room 23, Avant Building, Tallahassee, Florida 32304 (free to professional education persons and groups only)

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS Adult Education; Consumer Economics; *Consumer Education; Criterion Referenced Tests; Curriculum Development; Data Analysis; Data Collection; Economic Education; Educational Objectives; Elementary Secondary Education; *Evaluation; Group Activities; Higher Education; Inservice Teacher Education; Learning Activities; *Learning Modules; Measurement Instruments; *Performance Based Teacher Education; Pretests; *Program Evaluation; Skill Development; State Curriculum Guides; *Student Evaluation; Teacher Education; Teacher Role; Testing; Testing Problems; Test Validity

ABSTRACT

The document is the final in a series of four competency-based modules produced for use with inservice teacher education programs in consumer education. The main objective of the program is to assist curriculum developers and elementary and secondary classroom teachers as they plan, develop, implement, and evaluate a multidisciplinary consumer education program. This module presents numerous exercises to help teachers evaluate student performance in consumer education programs and offers suggestions for evaluating consumer education programs. The module also presents pretest exercises to assess knowledge of evaluation techniques, a glossary of evaluation terms, and suggestions for follow-up activities such as analyzing poor test items and constructing valid test items. Activities involve the teachers in matching and multiple choice exercises, taking true and false tests, writing essays, discussing evaluation techniques, giving examples of cognitive processes, recording observational data, and providing justification for various methods of evaluating student attainment. Background information is presented on norm referenced and criterion referenced tests, collecting and analyzing observational data, test data, measurement characteristics, and test validity. Upon completion of the exercises in this module, teachers are expected to be able to describe the end products, elements, and processes of evaluation; construct test items based upon a given content; and apply basic guidelines for writing objective test items. (Author/DB)

Evaluation in Consumers' Education: Teacher Education Module 4.

Florida State Dept. of Education, Tallahassee.

Project Title:
COMPETENCY-BUILDING IN CONSUMERS' EDUCATION
THROUGH MULTI-DISCIPLINARY TEACHER TRAINING

This module was developed by the Florida Department of Education under a grant from the United States Office of Consumers' Education, Washington, D.C.
For further information about the material, contact:

Dr. Steve Woolard
Project Director and Social Studies Consultant
Florida Department of Education
Miles Johnson Building
Tallahassee, Florida 32304

or

Mrs. Kathleen Funderburk
Home Economics Consultant
Florida Department of Education
Knott Building
Tallahassee, Florida 32304

ACKNOWLEDGMENTS

These modules were field-tested with one hundred elementary and secondary teachers and curriculum supervisors in Florida. The following group of distinguished individuals served as a content Validation Panel for the modules: Norman Dodl, Professor of Education, Virginia Polytechnic Institute; Edna Johnson, Director of Consumer Services, Baltimore, Maryland; Lillian Mohr, Professor of Home Economics, Florida State University; Lee Richardson, Professor of Economics, Louisiana State University; and Stowell Symmes, Director of Curriculum, Joint Council on Economic Education. The final products reflect the many recommendations which were made by these persons.

The following individuals were principal consultants:

Module One —Dr. Joseph Hurst, Director of Consumer Education Project, University of Toledo

Module Two —Dr. James Charkins, Department of Economics, California State College, San Bernardino

Module Three—Dr. Dennis Hooker, Instructor, Brevard Community College

Module Four —Dr. Hazen Curtis, Emeritus Professor, Florida State University

Also, Dr. Frank Lowney, College of Education, Warren Wilson College, contributed to the modules.

The following persons served as the principal project staff: Dr. William Rader, Dr. Stanley Kmet, Mrs. Kathleen Funderburk. Dr. Pat Spears was co-director at the onset of the project.

TABLE OF CONTENTS

ABOUT THE MODULES

This module is one of four competency-based modules produced for use with in-service teacher education programs in the area of consumers' education. The principal purpose of the modules is to assist curriculum leaders and teachers with the planning, development, implementation and evaluation of a multi-disciplinary consumers' education school program. While the modules do not comprise a complete in-service program, they provide a substantial foundation for both elementary and secondary teachers.

Module 1 provides an orientation to this program area and presents a generalized curriculum framework for the shaping of instruction and curricula outcomes.

Module 2 presents a conceptual framework drawing on major consumer and economic concepts and economic generalizations which help comprise a core of competencies for the consumer educator. These concepts and generalizations provide the basis for curriculum building.

Module 3 provides a set of experiences designed to help teachers relate the above conceptual framework to existing traditional school subjects. Several exercises are provided to help the teacher integrate the above concepts and generalizations into an existing curriculum.

Module 4 provides experiences that will enable teachers to improve their techniques for the evaluation of student performance in consumers' education. Procedures for program evaluation are also outlined.

Each of the above modules is a self-contained product intended for teachers to complete at their own pace. One or all four modules can be used to improve competency in a given area. Each module contains a Terminal Performance Objective—a competency each teacher should acquire upon completion of the module. Enabling Objectives comprise subcompetencies teachers should achieve as they proceed through the module. Enabling Elements contain instructional content designed to help the teacher meet the Enabling Objective. Pre- and post-tests are provided to help the teacher determine whether the Enabling Objective has been mastered or where additional review is necessary. At the conclusion of the module a follow-up activity is suggested.

We sincerely hope that, beyond creating basic teacher competencies in Consumers' Education, these modules stimulate a strong interest and enthusiasm in this vitally important field.

RATIONALE, OBJECTIVES, AND HOW TO USE THIS MODULE

Rationale

To evaluate is to make meaningful decisions based upon the best information at hand. The thoughtful evaluation of the learning/teaching outcomes of consumers' education is the most effective way a teacher can assure that the efforts to instruct and the efforts to learn will produce the best results practically obtainable.

Evaluative decisions may be made during the course of the instruction and study of a unit, or at its beginning. The evaluations conducted at these times lead to decisions to continue the instruction and study as it is going, to change it in certain ways, or to start certain students on the study of new modules or units. Evaluations conducted to decide whether, and how best to continue a teaching/learning activity, or whether to start a new one are called formative evaluations.

Evaluative decisions may be made at the close of a teaching/learning sequence. The teacher and the students need to know what has been learned during the course of the unit, and the teacher must decide upon its adequacy and quality. Evaluative decisions made at the close of a teaching/learning sequence for the purpose of summing up the adequacy and quality of the outcomes are known as summative evaluations. Evaluations upon which unit or course grades are based are summative evaluations.

The purpose of this module is to describe appropriate techniques for obtaining, in usable form, reasonably extensive, valid and dependable information about the principal aspects of student attainment; about the principal types of student input; and, about the principal types of teacher input which affect student attainment.

While it is recognized that in the making of evaluative decisions information pertaining to other fields (instructional materials, students' backgrounds, etc.) frequently will be needed, such topics are not included in this module. The objectives to which the content and activities of this module are directed are stated below.

Terminal Performance Objective

Given the information and exercises presented in this module, the teacher will demonstrate a mastery of its content by responding correctly to 70% of the test items included in a terminal test covering its enabling objectives.

Enabling Objectives

1. Describe evaluation in terms of (a) its end products, (b) its elements, and (c) its processes.
2. State the principal differences between test data and other observational data.
3. Describe each of five types of test items in terms of their respective forms.
4. Describe each of five types of test items in terms of the cognitive skills and processes that presumably will be used by students answering them.

5. Apply basic guidelines for the writing of all objective items.
6. Explain the meaning and the importance of test validity.
7. Explain the relationship between content and process validity.
8. Write items based upon a given content which will illustrate each of six cognitive skills or processes.
9. State the difference between norm and criterion referenced tests and illustrate uses of each reference system.
10. State valid reasons for considering observational data other than test scores when evaluating student attainment.
11. Describe a technique for recording such data in usable form.
12. Describe ways in which such data may be utilized when making evaluative decisions.
13. State professional reasons for documenting teacher input.
14. Use a form for documenting teacher input.

PRE-ASSESSMENT

Purpose

The purpose of the pre-assessment is to enable the reader to identify the sections of this module that he/she would benefit by studying. Each item in the assessment, or test, is tied to the section, or Enabling Element, of the module that presents the content tested by the item. By following the directions to be given, the pre-assessment will identify the Enabling Elements each person should study.

Directions

1. Prepare an answer sheet by writing the numbers from 1 through 20 in a column about one inch from the right edge of an ordinary sheet of paper. On this answer sheet, the answers can be written along the right edge of the page. They will be easy to score. (You may record your answers in the book, but scoring will be difficult.)

2. Read each item carefully, and read it completely.

3. For each item, there are a number of alternatives, or choices. Select the alternative that seems to be true, or to be the best one given. Write the letter number of that choice beside the number of that item on the answer sheet.

4. Turn the page, and work completely through the pre-assessment. Try all items, but if you cannot make a choice, leave the answer space blank and go on to the next item.

5. Turn the page and begin.

PRE-ASSESSMENT

Read each of the following statements carefully. Each is numbered. Some are TRUE; some are FALSE. Select the answer below which gives the numbers of ALL statements which are TRUE.

1. Any process by which data are collected is evaluation.
2. Evaluation conducted at the close of a sequence, or a period of instruction, is summative if it produces a decision about the value or merit of the outcomes of the activities over the sequence or period.
3. To be an evaluation, the decision must be based upon terms which can be clearly identified.
4. For formative evaluation, clearly stated terms and dependable data may be nice, but really are not necessary.
5. For good evaluation it is absolutely necessary to develop a comprehensive plan, and to carry out the plan.

1. In the paragraph above, the TRUE statements are those numbered:
   A. 1, 3, 5
   B. 2, 3, 5
   C. 3, 4; 5
   D. 2, 4, 5

Items 2 and 3 pertain to data obtained by administering tests and to data obtained through other types of observations. Read each statement carefully. Then mark the letter number of the alternative which is true of it.

2. The data usually can be checked by an independent observer.
   A. The statement tends to be more true of test data than of other observational data.
   B. The statement is equally true of data obtained from tests and from other types of observations.
   C. The statement tends to be more true of other observational data than of test data.
   D. The statement is true of neither.

3. The data may be affected to a considerable degree by personality factors in the observer.
   A. The statement tends to be more true of test data than of other observational data.
   B. The statement is equally true of data obtained from tests and from other types of observations.
   C. The statement tends to be more true of other observational data than of test data.
   D. The statement is true of neither.

Answer items 4 and 5 by matching each item type on the left with its description as often as you wish.

Item Type	Description
4. Alternative response	A. A statement or question followed by three or more possible answers, only one of which is correct or clearly the best
5. Multiple choice	B. A statement or question containing a blank into which the student is to write a word or phrase
	C. A question or statement of a task for which the student is to write his complete answer
	D. A question or a statement to which the student is to indicate his choice of two possible answers.

In the column on the right below, five cognitive skills and processes are listed and numbered. For the item type listed in question 6 and in question 7, select the alternative which records the numbers of the several cognitive skills and processes for which items of that type can be used effectively as measures.

Item Type	Cognitive Skills and Processes
6. Multiple choice: A. 1, 3, 5 B. 2, 4, 5 C. 1, 2, 3, 4 D. 5, 4, 3, 1	1. recall of information
	2. understanding
	3. opinion vs. fact
7. Essay: A. 1, 3, 4 B. 2, 4, 5 C. 3, 4 D. 5	4. sequencing
	5. organization and expression of personal reactions

The statement in each of the two following items may, or may not, be a good guideline for the writing of objective test items. Read each, and choose the response which best describes the statement as a guideline.

8. When wording item stems and the alternatives, use absolutes (always; never; all; none, etc.)

A. generally correct
B. not correct, because they generally signal an incorrect answer
C. generally okay, because their use makes item writing easier.

9. Errors students have made in the course of their study are the best source of erroneous alternatives for test items.
   A. false, because the students should not meet their errors on test
   B. whether true or not, it is unethical to use them because their use will make the items tricky
   C. correct as stated.

10. The validity of a test is its most important property.
   A. false, because the processes it actually measures are more important
   B. true, no other quality can make up for poor validity, but good validity can compensate for other possible weaknesses
   C. false, because if it is not valid for the purpose written, it still may produce interesting results
   D. true, because validity is a general property independent of the purposes of the teacher and characteristics of the students.

Read each of the numbered statements very carefully and then select the alternative below which lists the numbers of all that are TRUE.

Content and Process Validity

1. If the items in a test deal with the content actually covered in a unit or module, the test possesses content validity.
2. If the items require the students to employ the cognitive processes utilized in their study of the content, the test possesses process validity.
3. Content can be tested without involving one or more cognitive processes.
4. Each cognitive process must be tied to some element of content because each process must operate upon some content; that is, processes cannot operate in the abstract.
5. Any cognitive process can be tested by a single item if the item deals with suitable content.

11. I judge the TRUE statements to be:
   A. 1, 4, 5
   B. 2, 3, 5
   C. 1, 4
   D. 1, 2, 4

For each of the next two items, select the letter number from the list below of the cognitive skill or process that students probably would use in responding to it. You may use any skill or process as often as you wish.

Skill or Process

A. generalizing
B. inference
C. classification
D. hypothesis development

12. A girl found that she could invest in each of the following types of securities at the interest rates shown below:

U.S. savings bonds – 3.75%
national bank savings account – 4.5%
savings and loan account – 5.0%
first mortgage on a home – 6.5%
second mortgage on a home – 8.75%

In one short sentence, what probably accounts for the increase in interest rate as the types of investment change as you read down through this list?

Skill or process numbered _____ probably would be used by student responding to this item.

13. A number of fibers are listed below. Write the letter before each indicating whether it is:

A. an animal fiber
B. a vegetable fiber
C. a synthetic fiber

_____ linen
_____ dacron
_____ wool
_____ cotton
_____ nylon

Skill or process numbered _____ probably would be used by students responding to this item.

14. A teacher stated that two students failed to answer correctly 80% of the questions on the test on the module they had just completed, and that these two students therefore could not undertake the study of the next module.

Obviously the teacher was interpreting the test results in:

A. a norm referenced manner
B. an erroneous manner
C. a criterion referenced manner
D. an unethical manner

15. Which of the following are the two most important reasons for considering observational data other than test scores when evaluating student attainment?

Most Important Reasons

A. 1, 2

B. 1, 3

C. 2, 3

D. 3, 4

Possible Reasons

1. Certain objectives can be assessed only in terms of the behaviors of students in the course of their activities.
2. Some students do not do well on tests.
3. Some things students do probably indicate that changes in their learning are taking place before the results show up on test scores.
4. Observations of other behaviors give the student a defensible basis for awarding better grades to some students and lower grades to other students.

16. The following paragraph contains statements describing a technique for recording observational data other than test scores. Not all of the statements may be correct. Read the paragraph carefully and choose the alternative which lists the numbers of the statements which are correct.

The numbers of the correct statements:

A. 1, 2, 4, 6

B. 2, 3, 4

C. 3, 4, 5, 6

D. 2, 5

Description

(1) Observations of unscheduled student behavior should be recorded as anecdotes, primarily for use by the teacher making the observation. (2) Each record should state just what the student did and contain a brief notation of the circumstances under which the behavior occurred. (3) Each entry also should contain the teacher's inferences about the students' motives and the affective effects upon other students. (4) Prior to making the actual observations, the teacher should decide just what behaviors she will look for, and list them, keeping the list brief. (5) The behaviors to be observed may include any that interest the teacher, because she has no way of knowing to what she will relate them. (6) A simple notebook page for each pupil, listing his name, providing a column for dates, and blank space for recording the observed behaviors probably is the best record system to start with.

17. Listed on the right are possible ways of using observational data. Choose from the alternatives on the left the one that lists the generally advisable ways of using observational data.

Generally advisable ways to use other observational data

A. 1, 2, 3, 4, 5

B. 2, 3, 4, 5

C. 2, 3, 4

D. 1, 5

Possible Ways

1. Quantify them and treat the results mathematically.
2. Plan how best to use the records for all students if they will be used in arriving at the grade for any student.
3. Use the records for counseling individual students and in conference with individual student's parents.
4. Use them as one part of a student's record, being aware that the behaviors recorded constitute only a part of the total of any student's behaviors.
5. Use them in a completely unstructured and unplanned manner so you will not prejudice the behaviors of any student.

18. Which of the following are valid professional reasons for documenting teacher input into the learning/teaching process?

Valid Reasons

A. 1, 2, 3

B. 1, 2

C. 2, 3

D. 1, 3

Possible Reasons

1. Documenting the several things the teacher does makes them evident to the teacher.
2. Documenting the several things the teacher does provides evidence that the teacher did do something.
3. Documenting the several things a teacher does enables her to redistribute her effort and thus become more effective.

Classify the specific teacher activity given in each of the two following questions by recording the letter number of the type of teacher input to which it belongs from the list on the right.

Specific Teacher Activity

19. Wrote items for end-of-unit test (1 hour)

Type of Teacher Input

A. Planning and Preparation: Strategies
B. Other Activities: Independent Study
C. Direct Instruction: Large Group
D. Direct Instruction: Preparing Tests

20. Prepared presentation on Consumers' Education for parents' club (2 hours)	A. Related Activities: Committee B. Related Activities: Community C. Planning: Assemble Instructional Materials D. Direct Instruction: Large Group

INSTRUCTIONS

You may check back over any items if you wish.

Turn the page and check your answers with those given in the Answer Key.

DIAGNOSTIC ASSESSMENT

ANSWER KEY/ENABLING ELEMENT RELATIONSHIPS

Directions: Check your answers against the correct answers given in column two below. Record as your score the total number of items you answered correctly.

Then for each item answered incorrectly, study the corresponding enabling element given below. It is suggested that you study the enabling elements in numerical order (for example, enabling element numbered 5 should be studied before enabling element numbered 7).

Diagnostic Assessment Item Number	Correct Answer	Enabling Element Number	Go To Page
1	B	1	13
2	A	2	18
3	C	2	18
4	D	3	21
5	A	3	21
6	C	4	27
7	D	4	27
8	B	5	31
9	C	5	31
10	B	6	35
11	D	7	37
12	A	8	42
13	C	8	42
14	C	9	52
15	B	10	54
16	A	11	55
17	C	12	59
18	D	13	62
19	D	14	63
20	B	14	63

GLOSSARY

Content—the information, subject matter, ideas, etc., presented during instruction and presumably to be tested.

Enabling Element—the information printed in a module to enable the reader to attain the objective to which it is related.

Enabling Objective—one of the quite specific, smaller objectives, the attainment of which will contribute to the mastery of the larger terminal objective.

Generalization—a general statement linking two or more concepts; a principle.

Inference—an explanation an individual develops to explain, or account for, the occurrence of a relatively complex event.

Item Alternative—one of the choices offered as a possible answer or correct response to an item.

Item Stem—the basic statement, proposition, or question with which a test item deals.

Observational Data—records of events or of performances actually observed or seen; contrasts with hearsay records, beliefs, "hunches," etc.

Process—manner in which content, as defined above, is handled (manipulated) in the mind. A process is a way of mentally handling one or more ideas.

Reliability—the dependability of test results in the sense that if the same measures were given to the same students again under the same conditions, the results would be the same.

Supply Type Items—items that offer students no alternatives or choices of answers. Typically they contain blanks in which the student must write in, or supply, the correct word or phrase.

Terminal Performance Objective—the thing the student is to be able to do upon the completion of the module presented. In this module, the terminal performance objective is to answer correctly 70% of the items on the Post-Assessment.

Validity—the extent to which a test measures precisely what it was written to measure, and measures nothing else.

ENABLING ELEMENT 1

Evaluation in Consumers' Education

Evaluation is a process which leads to a decision as its end product. The decision may pertain to the outcomes of a past course of action, or activity, and be a statement of the extent or the value of the outcomes. For example, at the close of a reporting period, the teacher reviews the record of the work done by a student, decides its value, and awards the grade which best expresses the decision made. This type of decision is a summative evaluation.

The end product of an evaluation may be a decision made for the purpose of formulating new or modified lines of action, or strategies, designed to improve future outcomes over those realized in the immediate past. This type of decision is a formative evaluation. A teacher, on the basis of a short check test and observation of the students' study behavior, may identify a half dozen students who are not achieving satisfactorily, and who appear to be stymied in their study efforts. The teacher evaluates their situation, that is, makes a decision, for the purpose of enabling them to improve their future achievement. The types of decisions may vary, and the purposes for making them may vary, but a process must produce a decision if it properly is to be called an evaluation.

But all decisions are not the end products of an evaluative process. We all make impetuous decisions, and are unable to explain how we arrived at them. Such are not based upon an evaluation. An evaluation results in a decision for which we can state the elements that entered into its making. By the elements we mean the components, factors, or considerations upon which it rests. For example, a clear statement that students' grades will be based upon their test scores, their written class work, their participation in class discussion, and observed indications of changes in their values and attitudes affords a basis for making decisions that properly may be designated summative evaluations.

To utilize the stated components, factors or considerations in the making of evaluative decisions, teachers must have an adequate amount of valid and dependable information, data, about each. To obtain such data, plans must be instituted for its collection and for its processing into forms that are usable in the making of the decisions to be made. For example, if the students' term grades will be based, in part, upon the average (or median) of the percentages of the questions they answer correctly on each of three tests, the tests must be planned, the items written, the tests administered, the answers scored and the scores processed to yield data in usable form.

In summary, an evaluation is a decision pertaining to a clearly stated subject; a decision for which we can identify the components, factors, or considerations that entered into its making; and a decision for which we possess and utilize valid and dependable data pertaining to each component, factor, or consideration that entered into its making.

In terms of the proeesses involved, it is necessary to develop a plan which states just what is to be evaluated; states the components, factors, or considerations that

will enter into the making of the decision or decisions; specifies the data needed for each component, factor or consideration; and indicates the means by which each type of data will be obtained and processed into usable form; carry out the plan for data collection and processing, and utilizing the data obtained, make the decision or decisions necessary to evaluate the product or process stated.

Most teachers do not have much difficulty in determining just what they wish to evaluate. With some thought, they typically can determine the components, factors, and considerations that should enter into the making of their decisions about the outcomes or processes they decide to evaluate. Because obtaining an adequate amount of valid and dependable data pertaining to each component, factor, or consideration, and processing the data into usable forms are the recurring problems, these are topics addressed in much of the remainder of this module.

PRACTICE EXERCISE 1

In another module you were presented the **concept cluster: Production and Consumption of Goods and Services.** It dealt specifically with the concepts:

A. Good
B. Service
C. Production and Consumption
D. Household
E. Capital Good
F. Distribution

Assume that you will teach this concept cluster to pupils in the fifth grade, and that you will evaluate the effectiveness of the learning/teaching activities.

In the exercise below, three essentials of evaluation are listed in the column on the left. For each essential in the first column, record the letter number of the description on the right that best fits it.

Essentials of Evaluation in Consumers' Education	Description of Possible Essentials
_____ 1. A decision pertaining to a clearly stated subject	A. I shall base my decision upon scores on an 18 item test made up of three items pertaining to each concept. Correct responses to any two of the three items pertaining to each concept will be accepted as evidence of mastery of that concept.
_____ 2. Clearly identified components, factors, or considerations that will enter into the evaluation	B. Which of the students can proceed to the next concept cluster on the basis of their having mastered five of the six concepts presented in this cluster?
_____ 3. Valid and dependable data to be utilized when making the evaluation	C. Students are so different, and I know my students so well that I will just decide which ones need further study before going on to the next concept cluster.
	D. Following each student's name, I shall record in order the number of questions pertaining to each concept answered correctly, and then count the number of "2's" and "3's" for each student to determine how many of the six concepts were mastered.

Check your answers with the key on page 17.

PRACTICE EXERCISE 2

The column on the left lists five Teacher's Activities and the column on the right lists the three principal Evaluation Processes. Match each activity with its process by recording the letter number of the process in the space preceding each activity. Use any process as often as you wish.

Teacher's Activity

_____ 1. Administered a test on the concepts in the cluster entitled "Production and Consumption of Goods and Services"

_____ 2. Told the class that whether or not a student may go on to the next concept cluster will depend upon mastery of five of the six concepts in this cluster.

_____ 3. Scored the test in terms of the number of questions pertaining to each concept answered correctly by each student.

_____ 4. Counted the number of concepts for which each student individually answered correctly 2 or 3 questions, and if the total was 5 or more, placed him in the group to study the next cluster.

_____ 5. Mastery will be determined by responding correctly to two out of three questions pertaining to each concept.

Evaluation Process

A. developing a plan for the total evaluation

B. carrying out the plan for data collection

C. utilizing the data in making the evaluative decisions

Check your answers with the key on page 17.

Key for Practice Exercise 1

1. B
2. A
3. D

Key for Practice Exercise 2

1. B
2. A
3. B
4. C
5. A

ENABLING ELEMENT 2

Similarities and Differences Between Test Data and Other Observational Data Pertaining to Student Attainment

Student attainment enters into most evaluations in which teachers are involved. Some aspects of attainment, such as knowledge gained, skills mastered, attitudes developed, etc., may be the subject of the evaluation, or they may be the elements upon which the effectiveness of something else (teaching strategy, for instance), is to be evaluated. Data pertaining to student attainment may be divided into two broad types: test data and other observational data.

Basically, all data are records of observations made by somebody or reported to have been made by somebody. All are records of real or purported observations. The datum that the time of sunrise on a given day was 6:20 a.m. is one whose accuracy is subject to independent verification. The datum that a ghost appeared at a certain spot at a certain time may be a record of a reported observation, but it is not subject to independent verification, and hence its accuracy is subject to question. A dozen people standing near each other, all having synchronized chronometers, probably would have differed in their reported times of sunrise by only a few seconds. The reported time was reliable within a small margin of error. The reports of a dozen people near the spot where the ghost allegedly appeared doubtless would vary widely among themselves. Hence the reported sighting would have low reliability. Sighting the sunrise is a valid way of determining its time of rising. Sighting a ghost may be a valid way of establishing the reality of ghosts, but the accuracy and reliability of such data are so poor that the validity of the reported sighting well may be questioned. While all data used in evaluating educational outcomes and processes should originate as observations, in practice they vary greatly in terms of their accuracy, reliability, and validity.

Tests yield records of the actual performance of the student taking the test. The records have a certain permanence, and are there for anyone to see and to examine. This much is true of all types of short answer tests, of essay tests, and of tests requiring the production of a durable product like a picture or a piece of furniture.

Other observations may deal with transitory behaviors, lacking any degree of permanence and leaving no record in and of themselves. A nod of agreement with an expression of a value; a student obviously comparing the advertisements for two garments; selecting an orange instead of pie at the cafeteria counter may be observed only when and while they are happening, and in and of themselves they leave no record. They cannot be re-checked, and cannot be examined by anyone not present at the time of their occurrence.

Data derived from tests may be obtained whenever they are desired. The obtaining of test data is completely controllable by the teacher desiring to make that type of observation. Observations of other behaviors generally must be made whenever they occur, and their occurrence is unscheduled and generally unpre-

dictable and uncontrollable. Under such circumstances, the teacher misses observing the occurrences of many behaviors, and hence the data pertaining to them generally understate their frequency by some unknown amount.

Test data practically are free of personality differences which may greatly affect those derived from other observations. All students make the same types of records on their answer sheets, and these records are equally observable by all teachers. The observable manifestations of affective responses vary greatly from student to student, being fairly obvious for some, and practically undetectible for others. Teachers, too, vary in their ability to detect such responses. Some are quite perceptive, and others simply are not. All probably are selective in the detection of affective responses. Each person is more likely to detect responses similar to those he makes, than to detect responses which are not. Student differences and teacher differences combine to lower the reliability and even the validity of much observational data.

Not all test data are equally objective, valid, and reliable. Generally data derived from so-called objective tests in which the student marks one correct answer are quite dependable. Data derived from essay tests generally are somewhat less dependable because of scoring problems. The students' answers vary greatly in accuracy, completeness, style and legibility. Determining precisely how much each student knows is difficult and the teacher's standards may shift during the process of scoring the papers. Deriving accurate data from objective tests is relatively easy; deriving accurate data from essay tests is more difficult and the dependability of the scores derived often is affected adversely.

Data derived from tests generally originate in numerical terms, being the number of questions answered correctly or the number of points made on an essay test. They may be manipulated by using common arithmetical processes. Thus the raw data may be converted into percentages, into some standard scale values, their averages computed, ranked, and grouped as the teacher may wish. Each step in the process from scoring to final reporting may be checked for accuracy.

Data derived from most other types of observations originate in verbal terms. They report the teacher's observations. Therefore, the data are subjective reports whose precise meaning may not be clear to another reader, or even to the writer after a period of time. Being verbal reports, their processing and synthesis is difficult and not exact. Typically, the data can only be reviewed, and subjective judgments made about the frequency and quality of the behavior described. If attempts are made to convert the data into numerical terms, the resulting data retain all the weaknesses of the original verbal reports. Additionally, the data assume the weaknesses of nominal or coarse ordinal measures with which little properly can be done, arithmetically.

PRACTICE EXERCISE 3

Check the correct column, "true" or "false."

Compared to test data, observations of other types of student behavior are:

T	F		
_____	_____	1.	more permanent
_____	_____	2.	more readily confirmed by other observers
_____	_____	3.	less easily interpreted
_____	_____	4.	more accurately treatable by mathematical processes
_____	_____	5.	less reliable
_____	_____	6.	more objective
_____	_____	7.	less complete

Check your answers below.

(Answers: 1-F; 2-F; 3-T; 4-F; 5-T; 6-F; 7-T)

ENABLING ELEMENT 3

Formal Characteristics of Each of Five Item Types

Figure 1, which follows, supplies in its second column a brief description of the form of each of the five most commonly used types of items, namely: multiple choice, matching, alternative response, supply, and essay. In Figure 1, the form of test items will be described only in terms of their essential elements, or parts that must be presented to compose an item of the type designated.

For each type of test item listed in column 1, read carefully the description of its form in column 2. Then try Practice Exercise 4, which follows.

Figure 1. Characteristics of five types of items in terms of form and of processes measured.

Item Type	Form	Cognitive Processes
1. Multiple Choice	Statement, or stem, plus three or more choices, one, and only one of which will correctly complete the statement or answer the question presented in the stem. Typically, the stems are numbered as questions, and the choices are lettered a, b, c, ---.	Efficiently measures recall of knowledge of factual information. Can also be constructed as an effective measure of understandings, thinking skills, comparison, contrast, sequencing, opinion vs. fact, etc. Not recommended to measure organization and expression of personal thoughts, reactions, etc.
2. Matching	Stem directs student to select from the entries in one column the one or more which are associated with each object, process, or attribute listed in the other of the two columns immediately following.	Efficiently measures recall of knowledge of factual material. Can also be constructed to measure effectively understandings and selected thinking skills, comparison, contrast, sequencing, opinion vs. fact, etc. Not recommended to measure organization and expression of personal thoughts, reactions, etc.
3. Alternative response	Statement or stem states a proposition to which alternatives offer two reasons that are opposite in meaning; often true-false; fact-opinion; greater than-less than, etc. Items generally are brief.	Efficiently measures knowledge of things or events that naturally can exist in only one way or in the opposite way, or that can be represented as so existing. Has very limited value as means of measuring higher order cognitive processes without making the stem complex and involved. Cannot be used to measure ability to generate, compose, and express students' personal thoughts, reactions, etc.

4. Supply	Statement or stem requires the supplying of a critical word, phrase, or short clause at the point indicated to complete it, making it true. Different words or expressions may be used provided their meaning does not differ from that of the expected response.	Efficiently measures recall of knowledge of factual information. May be used to measure other cognitive processes and skills, but to do so, the stem may become so complex that serious reading problems result. Not recommended to measure organization and expression of personal thoughts, reactions, etc.
5. Essay	Statement or question clearly specifying the subject and the manner and terms in which the student is to respond to it. Sufficient blank space should be provided immediately following the stem for the student to write the response.	Not recommended for the recall of factual information and, in most instances, not for the measurement of understandings and several other cognitive skills and processes because other item types may be used to do so more efficiently. Effective measurement of ability to organize and express personal thoughts and reactions, and to weigh values and alternative arguments.

PRACTICE EXERCISE 4

For each item below, indicate its formal type by recording the letter of the correct answer.

A. Multiple Choice
B. Matching
C. Alternative Response
D. Supply
E. Essay

Type

_____ 1. A beautician produces a good
A. True
B. False

_____ 2. For each statement on the left, record the letter number of the concept in the right column with which it belongs.

Statement	**Concept**
_____ 1. washing cars	A. Good
_____ 2. grower of potatoes	B. Service
_____ 3. an auto tire	C. Producer
_____ 4. a potato	

_____ 3. In a short paragraph, explain the concept "distribution."

_____ 4. Goods which are produced for the purpose of producing other goods are called _____ goods.

Check your answers below.

Answers: 1-C; 2-B; 3-E; 4-D

POST-ASSESSMENT

Read each of the following statements carefully. Each is numbered. Some are TRUE; some are FALSE. Select the answer below which gives the numbers of ALL statements which are TRUE.

1. Any process by which data are collected is evaluation. 2. Evaluation conducted at the close of a sequence, or a period of instruction, is summative if it produces a decision about the value or merit of the outcomes of the activities over the sequence or period. 3. To be an evaluation, the decision must be based upon terms which can be clearly identified. 4. For formative evaluation, clearly stated terms and dependable data may be nice, but really are not necessary. 5. For good evaluation it is absolutely necessary to develop a comprehensive plan, and to carry out the plan.

1. In the paragraph above, the TRUE statements are those numbered:

   A. 1, 3, 5
   B. 2, 3, 5
   C. 3, 4, 5
   D. 2, 4, 5

Items 2 and 3 pertain to data obtained by administering tests and to data obtained through other types of observations. Read each statement carefully. Then mark the letter number of the alternative which is true of it.

2. The data usually can be checked by an independent observer.

   A. The statement tends to be more true of test data than of other observational data.
   B. The statement is equally true of data obtained from tests and from other types of observations.
   C. The statement tends to be more true of other observational data than of test data.
   D. The statement is true of neither.

3. The data may be affected to a considerable degree by personality factors in the observer.

   A. The statement tends to be more true of test data than of other observational data.
   B. The statement is equally true of data obtained from tests and from other types of observations.
   C. The statement tends to be more true of other observational data than of test data.
   D. The statement is true of neither.

ENABLING ELEMENT 4

This Resource has two parts. Work through both parts before attempting Practice Exercise 5.

Measurement Characteristics of Each of Five Item Types

The types of cognitive skills and processes which each of the item types may be used to measure are reported in Column 3 of Figure 1. Review Figure 1 and note that several of the item types may be used to measure a wide range of skills and processes; the same skill may be measured by using any one of several item types; alternative response items and supply items are limited with respect to the skills which they can measure; and essay items should be used to measure only objectives that require the student to put together an original response.

There are other cognitive skills and processes which the teacher may wish to measure. One or more of these item types may be used to measure most of the skills which teachers wish to measure. It may require some ingenuity to write an item of a type which will measure a particular skill. However, with thought and experience, the typical teacher can write objective items which measure most of the cognitive skills and processes developed in the study of Consumers' Education.

Problems Inherent in the Measurement of Cognitive Skills and Processes

A word of caution may be in order. A cognitive skill or a cognitive process is an operation which we assume goes on in a student's brain and nervous system. The fact is that we cannot observe what goes on in a student's brain and nervous system, and the further fact is that his report of what went on often is sketchy or basically inaccurate.

We may write an item to measure the ability to utilize a specific skill, but there is no way of knowing what skill actually is employed by the student in responding, even if the answer is marked correctly. It is intellectually economical to use the simplest skill the respondent finds possible, and good practice to do it. This principle leads to the perverse situation in which when responding to a given item the good student who has a large fund of knowledge will respond by recall, and the poor student with a meager fund of knowledge will employ inference, comparison, or some higher skill in arriving at his answer. As a result, the teacher may accurately conclude that 75% of the students were able to answer the item correctly, but the inference that 75% can perform the skill the item was designed to measure may be tenuous.

ENABLING ELEMENT 4

This Resource has two parts. Work through both parts before attempting Practice Exercise 5.

Measurement Characteristics of Each of Five Item Types

The types of cognitive skills and processes which each of the item types may be used to measure are reported in Column 3 of Figure 1. Review Figure 1 and note that several of the item types may be used to measure a wide range of skills and processes; the same skill may be measured by using any one of several item types; alternative response items and supply items are limited with respect to the skills which they can measure; and essay items should be used to measure only objectives that require the student to put together an original response.

There are other cognitive skills and processes which the teacher may wish to measure. One or more of these item types may be used to measure most of the skills which teachers wish to measure. It may require some ingenuity to write an item of a type which will measure a particular skill. However, with thought and experience, the typical teacher can write objective items which measure most of the cognitive skills and processes developed in the study of Consumers' Education.

Problems Inherent in the Measurement of Cognitive Skills and Processes

A word of caution may be in order. A cognitive skill or a cognitive process is an operation which we assume goes on in a student's brain and nervous system. The fact is that we cannot observe what goes on in a student's brain and nervous system, and the further fact is that his report of what went on often is sketchy or basically inaccurate.

We may write an item to measure the ability to utilize a specific skill, but there is no way of knowing what skill actually is employed by the student in responding, even if the answer is marked correctly. It is intellectually economical to use the simplest skill the respondent finds possible, and good practice to do it. This principle leads to the perverse situation in which when responding to a given item the good student who has a large fund of knowledge will respond by recall, and the poor student with a meager fund of knowledge will employ inference, comparison, or some higher skill in arriving at his answer. As a result, the teacher may accurately conclude that 75% of the students were able to answer the item correctly, but the inference that 75% can perform the skill the item was designed to measure may be tenuous.

PRACTICE EXERCISE 5

For each of the four item types listed in the first column, check the letter numbers of all skills or processes for which that item type may be used as an efficient measure.

Item Type

1. **Supply**
   - _____ A
   - _____ B
   - _____ C
   - _____ D
   - _____ E
   - _____ F
2. **Multiple choice**
   - _____ A
   - _____ B
   - _____ C
   - _____ D
   - _____ E
   - _____ F
3. **Alternative response**
   - _____ A
   - _____ B
   - _____ C
   - _____ D
   - _____ E
   - _____ F
4. **Essay**
   - _____ A
   - _____ B
   - _____ C
   - _____ D
   - _____ E
   - _____ F

Skill or Process

A. recall of facts

B. understanding

C. comparison

D. organization of personal thoughts

E. true or false

F. weighing values

5. If a teacher writes a very good item to measure the ability to perform a certain cognitive process and a student marks the correct answer, which of the following statements may be true? Check all that *may* be true.

_____ A. The student used the process intended

_____ B. The student used a simpler process

_____ C. The student used a more complex process

_____ D. The teacher really does not know which process the student used.

Check your answers with key on page 30.

Key for Practice Exercise 5

1. Supply

   X A

2. Multiple Choice

   X A
   X B
   X C
   X F

3. Alternative Response

   X A

4. Essay

   X D
   X F

5. Multiple Choice Question

   X A
   X B
   X C
   X D

ENABLING ELEMENT 5

Basic Guidelines for the Writing of Objective Items

Writing objective test items is an art and not a science. Every experienced item writer develops individual guidelines, some of which are clearly formulated, and others which are followed intuitively. Inexperienced writers are advised to formulate and to revise guidelines as they gain experience.

The following list of guidelines is not exhaustive. They are intended to guide the efforts of relatively inexperienced writers and to be modified and supplemented by writers as they gain experience.

For convenience, the guidelines are grouped under general headings. The order of the groupings does not mean that one is more important than others, because to disregard any of them will spoil the items written.

Writing and Organization

Employ as simple language and sentence structure as the subject matter and the process to be measured permit. The task of reading the question should not interfere with the student's effort to answer it.

Eliminate all ambiguities in the stem and the alternatives.

"Oranges, roses, and pinks are:

a. colors
b. plants
c. fruits," is ambiguous because the items named may be either colors or plants.

Avoid absolutes in wording the item and in the alternatives. "Always, never, all, none; every, no, etc.," generally signal a false statement or erroneous alternative.

Organize the presentation of the stem and the alternatives so the student finds the possible answers preferably along the left margin of the page. The student should not spend time locating the physical position of the elements he needs. Placing the alternatives to which the responses are to be attached along the left margin makes hand scoring much easier for the teacher.

Item Stems

The stem should state clearly the terms in which the student is to respond to the stated or implied question.

"Which of the following investments is best in terms of security of the capital invested?" is a very different item from *"which of the following investments is best in terms of return on the capital invested?" "Which of the following investments is best?"* is impossible to answer.

Except in the case of supply type items, the stem plus the alternatives should provide all the elements of information required to respond to the item.

"The simple interest on $100 for one year is _____." and *"The interest for one year at 6% on $100 is _____,"* are both deficient. The first omits the rate. The second probably would be answered one way by the poorer students who know only simple interest, and another way by better students who may know something about compound interest if one of the distractors is somewhat greater than $6.00.

Alternatives

The best source of erroneous alternatives is the errors students made in the course of their study of the topic tested.

Generally the correct alternative is easy to write. Students' errors are, or were, plausible to them. By using their errors as alternatives, the item becomes an observation of the extent to which errors have been corrected through the instruction and study.

For each item there must be one, and *only* one correct answer, or the instructions must state that the student is to mark all that are correct.

The situation sometimes arises in which the general instruction has pointed to one answer as correct, but on the basis of more extensive information another answer properly is judged to be correct.

"The savings accounts in all savings and loan associations are insured up to a certain amount" probably would be marked true by students studying only about federal s&l's, but false by those knowing that some states charter state s&l's without requiring deposit insurance.

PRACTICE EXERCISE 6

Directions: There are five sample items in this practice. Below each item is a list of alternatives. In the blank provided, check the alternative which describes the sample item's greatest weakness.

1. *A woeful deficiency in any family's department of its treasury is precipitated when its gross expenditures exceed by large sums the net income it receives.*

_____ *True*

_____ *False*

The greatest weakness of the item as stated is:

_____ A. It is not really a "true/false" question.

_____ B. The language and structure are unnecessarily difficult.

_____ C. The item is ambiguous.

_____ D. Not enough information is given in the item.

2. *The best credit card is:*

_____ *A. American Express*

_____ *B. local store*

_____ *C. BankAmericard*

_____ *D. Gasoline company*

The greatest weakness of the item as stated is:

_____ A. The language is too difficult.

_____ B. The arrangement of the alternatives is confusion.

_____ C. There is a poor choice of alternatives.

_____ D. The terms in which the student should judge are not given.

3. *Interest, wages, and medical services of a family are:*

_____ *A. Sources of income*

_____ *B. Costs*

_____ *C. Savings*

The greatest weakness of the item as stated is:

_____ A. The item is ambiguous; for some families the items mentioned may be sources of income, and for other families they may be costs

_____ B. The alternatives represent ideas students would not ordinarily have

_____ C. It uses absolutes which make the correct answer clear

_____ D. The organization of the item is confusion.

4. *If a purchaser pays cash, she can buy a clothes dryer for $160.00. If she charges it for 30 days, the cost will be $5.00 more. What is the true annual percentage rate of interest if she charges?*

_____ *A. 3%*

_____ *B. 11½%*

_____ *C. 37½%*

_____ *D. 25%*

The greatest weakness of the item as stated is:

_____ A. more than one answer may be correct

_____ B. the correct answer is not given

_____ C. the placement of the alternatives is confusing

_____ D. not sufficient information is given to enable the student to answer the question.

5. *All items always cost more if they are charged.*

_____ *True*

_____ *False*

The greatest weaknesses of the item as stated is:

_____ A. the sentence structure is poor

_____ B. the use of the two absolutes, "all" and "always" suggest strongly that the statement is false

_____ C. both answers may be correct

_____ D. the stem does not supply all the information needed to respond to the item.

Check your answers with the key below.

Key for Practice Exercise 6: 1-B; 2-D; 3-A; 4-C; 5-B.

ENABLING ELEMENT 6

The Meaning and Importance of Test Validity

The validity of a test is the extent to which it actually measures what it was constructed to measure. If a test was constructed to measure knowledge of terminology used on labels of cans of food and the students respond purely on the basis of their knowledge of that terminology, the test is perfectly valid. If their responses are based in part upon their ability to read the items, then the validity of the test is reduced to the extent that reading competence influenced the responses actually made.

The term validity often is used loosely. This loose usage causes and sustains misconceptions which hinder constructive understandings of test validity. The first misconception is that a general property called "validity" exists and can and should be inherent in good tests. Validity as a general property is a fiction, indeed a delusion. Validity is very specific.

The second misconception is implied when even specific validity is spoken of as an "all-or-none proposition." Items and tests are valid measures of some specific thing to some extent, usually greater than zero and less than 100%. The question about a test of check writing skills should be "how valid is it?" not "is it, or is it not, valid?" In practice, an item or a test should be just as valid as the writer can make it for the purpose for which it was written. Almost certainly its validity will be greater than zero, and almost certainly it will be less than 100%.

For classroom tests, the actual validity of an item or a test also is influenced by the abilities of the students taking the test and by their prior instruction. A given item or test may be quite valid as a measure of check writing skills when given to middle school students who can read the items easily. It may possess very low validity as such measure if given to fourth graders whose reading problems block their efforts to cope with the items. In this situation, it may operate as a quite valid test of reading comprehension, but that is not what it was written to measure!

Another misconception seems to be that there is some way for a teacher to determine in advance just how valid the test is, or will be. Unfortunately, such is not the case. The best a teacher can do when preparing a test is to exercise the best judgment. The appropriate question to be asked is, "Will this item when given to these students lead the large majority of them to respond in terms of the competency I want it to measure?" If in the teacher's judgment it will, he/she has done the best to obtain operating validity.

The validity of a test is its most important property. If it actually does measure quite well what it was written to measure, it does the job. If it actually does not measure very well what it was written to measure, nothing else that it may do justifies its use for that purpose.

Validity is so important that the basic reason for giving attention to test planning, to item writing techniques and skills, to test format, to test administration, and to test reliability is to reduce to a minimum those things which might lower the operating validity of the test when used.

PRACTICE EXERCISE 7

Directions: For each item below, check the one best answer.

1. If a test measures exactly what it was written to measure and nothing else, it is:
   - _____ A. somewhat valid
   - _____ B. perfectly valid
   - _____ C. completely invalid

2. A really valid test is:
   - _____ A. valid for students of any age
   - _____ B. valid for any content
   - _____ C. valid for any cognitive process
   - _____ D. valid for certain students, and for a certain content and a certain process

3. A test written by a teacher generally is:
   - _____ A. valid to some extent, but not perfectly valid
   - _____ B. totally valid
   - _____ C. totally invalid
   - _____ D. not affected by validity

4. To prepare valid tests the teacher should:
   - _____ A. use a statistical formula
   - _____ B. disregard the abilities of the students
   - _____ C. use best judgment when writing each item
   - _____ D. disregard objectives when teaching the content or process tested

Check your answers with the key below.

Key to Practice Exercise 7: 1-B; 2-D; 3-A; 4-C.

ENABLING ELEMENT 7

The Relationship between Content and Process Validity

It is commonly emphasized that the classroom test should focus upon the content actually studied in the unit or module being tested. To the extent that it does, the test will possess content validity. To the extent that it does not, the test lacks content validity. The students probably will say that it is unfair.

Teachers have become concerned about the students attaining competence in utilizing certain cognitive processes and skills and have made such attainments major objectives of their instruction. Their tests should be valid measures of the cognitive processes which were objectives of their instruction.

The point that seems all too often to be missed in the discussions of testing content and testing processes is that content and process necessarily are two dimensions of every test item. The student must employ some cognitive process in responding to the content of each item, and no cognitive process can operate if it does not do so upon some content. The task of the test writer is to so bring content and process together in items that when responding, the student will utilize the desired skills or processes in handling the content of the items.

The diagram presented in Figure 2 represents the two dimensional nature of content and process and offers a technique by which the test writer can plan to tie elements of content to the cognitive processes he/she wishes students to utilize in formulating their responses.

Each row of the Figure is devoted to an element of the content presumably studied by the students for whom the test is to be written. If, in fact, each element has been studied, all items shown in its row will be valid in terms of content if they deal with that element.

Each column is devoted to a process. The entries show the number of items utilizing each process and the elements of the content to which they are to be tied. For instance, one or more recall items are to be written about all the elements except number two and number six. Classification items are to be written about two of the elements, etc.

The entry showing six items to be written which tie hypothesis development to the element "habits causing poor nutrition" illustrates two important principles in the testing of content/process. While most processes may be reduced to one or two steps, the more complex, of which hypothesis development is one, inherently require a number of sequential steps, and therefore require as many sequential items to measure them. Hypothesis development really is a six-step process and requires six items to completely cover it. The second principle is that each process must be tied to a content element large enough to accommodate its complexity. "Habits causing poor nutrition" is large enough to accommodate the complexities of hypothesis development; others may not be.

Figure 2. Test Planning in Terms of Content and Process

General Topic – Nutrition

Elements of Content	Processes				
	Recall	Classification	Inference	Generalizing	Hypothesis Development
1. Names of basic elements	1				
2. Typical foods		2		1	
3. Amount of intake	3			1	
4. Waste and destruction	2		1	1	
5. Relative Costs	2	2	1		
6. Habits causing poor nutrition			1	1	6
TOTAL	8	4	3	4	6

PRACTICE EXERCISE 8

Check each of the following statements which are *true*.

_____ 1. If a test deals with the content actually taught it possesses content validity.

_____ 2. Answering each question requires the exercise of some cognitive skill or process.

_____ 3. It is possible to write test items which measure content and require the student to use no cognitive skill or processes.

_____ 4. Items may be written which are pure measures of cognitive processes and skills and involve no content.

_____ 5. Every item must deal with some content and responding to it requires some cognitive skill or process.

_____ 6. Testing a sample skill or process requires only a small or specific bit of content.

_____ 7. Testing complex or involved cognitive skills and processes requires that the student be given, or that he already has, rather extensive blocks of content.

_____ 8. When planning a test, each element of content should be tied to the cognitive skills or processes to be tested with it.

Check your answers below.

Answers: X 1. X 2. ___ 3. ___ 4. X 5. X 6. X 7. X 8.

POST-ASSESSMENT

In the column on the right below, five cognitive skills and processes are listed and numbered. For the item type listed in question 1, select the alternative which records the numbers of the several cognitive skills and processes for which items of that type can be used effectively as measures.

Item Type

1. Alternative response:
   - A. 1, 3
   - B. 2, 4
   - C. 4, 5
   - D. 3, 5

Cognitive Skills and Processes

1. recall of information
2. understanding
3. opinion vs. fact
4. sequencing
5. organization and expression of personal reactions

The statement in each of the two following items may, or may not, be a good guideline for the writing of objective test items. Read each, and choose the response which best describes the statement as a guideline.

2. Except for supply type items, the stem plus the alternatives should omit at least one element of information required to respond to the item.
   - A. exactly reversed; not correct
   - B. correct as stated
   - C. even if correct, this would not be a guideline.

3. Errors students have made in the court of their study are the best source of erroneous alternatives for test items.
   - A. false, because the students should not meet their errors on test
   - B. whether true or not, it is unethical to use them because their use will make the items tricky
   - C. correct as stated.

Read the following paragraph very carefully. You will note that each sentence is numbered. After reading the paragraph, choose the alternative which lists the numbers of all sentences which are NOT correct.

1. A test is valid to the extent that it is reliable, that is, that it measures dependably what it does measure. 2. The validity of a test is very specific, that is, it may be valid as a measure of one thing and possibly totally invalid as a measure of other things. 3. Most tests that we construct are valid to some extent, but probably not 100% valid. 4. The same test will be equally valid as a measure of a given body of knowledge if it is administered to two groups of students differing significantly in

their command of basic skills. 5. The best a teacher can do when writing test items is to try to write each item so it will operate as a valid measure of what the teacher wants it to measure when given to the class for which it is written.

4. In the above paragraph, the sentences numbered _____ are NOT correct:

A. 1, 3, 5
B. 2, 3, 4
C. 3, 5
D. 1, 4

Read each of the numbered statements very carefully and then select the alternative which lists the numbers of all that are FALSE.

1. If the items in a test deal with the content actually covered in a unit or module, the test possesses content validity.

2. If the items require the students to employ the cognitive processes utilized in their study of the content, the test possesses process validity.

3. Content can be tested without involving one or more cognitive processes.

4. Each cognitive process must be tied to some element of content because each process must operate upon some content; that is, processes cannot operate in the abstract.

5. Any cognitive process can be tested by a single item if the item deals with suitable content.

5. I judge the FALSE statements to be:

A. 1, 2, 4
B. 1, 3, 4
C. 3, 5
D. 2, 4

Answers: 1-A; 2-A; 3-C; 4-D; 5-C

17

ENABLING ELEMENT 8

Examples of Content/Process Items

A. Recall

Recall is the process of calling into consciousness (remembering) information stored in memory. It is basic to all intellectual activity, and is economical of time and effort. Recall may operate with simple or complex, small or large segments of subject matter.

Examples

1. In terms of typical grocery store prices, from which of the following could a person most likely get the amount of protein needed for one day at the least cost?

   a. beef roast
   b. oranges
   c. dry beans
   d. margarine

2. Fill in the blank with a *word* that makes this statement true:

   A good is a _____.

B. Classification

Classification is the process of putting together things, actions, or events that belong together because of some basic, or general principle, characteristic, or tie.

Examples

1. Make an "X" before each of the following foods which properly may be classified as a carbohydrate.

   Foods

   _____ 1. bread
   _____ 2. butter
   _____ 3. bacon
   _____ 4. jelly
   _____ 5. potatoes
   _____ 6. cake
   _____ 7. red-eye gravy

2. Make an "X" before each item below which is a *service.*

   _____ A. washing a customer's car
   _____ B. an orange
   _____ C. an automobile
   _____ D. cutting a customer's hair
   _____ E. preparing a friend's tax return

C. Inference Making

Inferenee making is the process of deriving an explanation of the reason why things happened as they did happen.

Examples

1. A housewife began to:
   1. select cuts of meat with as little bone and fat as possible.
   2. buy perishable fruits in smaller quantities.
   3. select vegetables from open bins instead of prepackaged.
   4. serve food on children's plates.
   5. prepare casserole dishes from left-overs.

   I infer that her principal reason for doing all of these things was to:
   A. get better buys in meats.
   B. reduce waste.
   C. teach children to eat properly.
   D. avoid buying spoiled fruits and vegetables.

2. A lady showed her friends a new chair she had purchased. Its design, fabric, frame and finish looked like that of a $150.00 chair. There was no brand name on it. She said she paid about $80.00 and that the store had other chairs like it. She didn't tell where she bought it, but I infer that it probably was a:
   A. department store.
   B. discount store.
   C. "name brand" furniture store.
   D. general store.

D. Conceptualizing

Conceptualizing is the process of bringing seemingly unrelated and distinct things or events into a functioning whole. It is the process of developing a "big idea" into which a number of different things or events fit, and in terms of which all of the associated details can be handled cognitively.

The value of a concept demands upon the number of apparently separate things or events in an area it brings together. Typically, concepts "grow" with the study and experience of students. Simply because concepts are powerful tools used in the process of thinking, it is important in Consumers' Education that each of a few basic concepts "grow" as the students study and gain experience in the field.

In Consumers' Education, much of the testing of the process of conceptualizing will present items requiring the students correctly to relate new things or events to previously acquired basic concepts. To correctly relate new things or events to previously acquired basic concepts indicates that the basic concept has "grown" in the sense that the concept is larger and covers more things or events. Hence it is more useful and valuable.

Examples

(For convenience both examples will deal with the concept specialization.)

1. Below is a list of five stores. Beside each store are the things the store used to sell. The next list tells what the store **sells now.** Study the lists. Then put an X by the two stores that have become more SPECIALIZED.

Store	Used to Sell	Sells Now
_____ Bill's	orangeade and lemonade	orangeade
_____ Joe's	orangeade and lemonade	nothing
_____ Ken's	orangeade and lemonade	lemonade
_____ John's	orangeade and lemonade	orange and lemonade
_____ Ben's	orangeade and lemonade	lemonade and koolaide

2. Below are four countries and a description of their economies. Place an X by the two which need to increase in SPECIALIZATION.

Bongonia is heavily industralized with very little agricultural production. The country's major exports have been in the nature of heavy machinery, while imports have been consumer goods and agricultural products. A major world war has broken out and foreign trade has been completely cut off.

Amazonia is on a continent with many other countries. The other countries are all heavily industrialized, but have very inadequate agricultural production. Amazonia is self sufficient at a very low economic level. It has rich farm lands, but devotes 90% of its productive efforts to industrial production. This production is very inefficient due to the lack of adequate mineral resources.

Washingtonia is an industrial country which produces heavy machinery. It relies on the sale of the machinery to pay for its imports of general consumer goods. In recent years many other countries have begun competing with Washingtonia and the country is rapidly being forced out of the market.

Ecolonia is a large country which has enjoyed tremendous economic success over the past two hundred years. The standard of living has been high, due to profits from large exports of agricultural and industrial products and raw materials. Lately the gross national product has declined due to increasing shortages of minerals and wear on the land.

E. Hypothesizing

Hypothesizing, or the formulation of a tenable and testable prediction based upon the best available information is one of the more complex cognitive processes. As used under most circumstances in Consumers' Education, it requires that the

student first examine the information available in order to figure out what important relationships exist among the facts or data in hand. It then requires that the student reason that if these relationships exist, making a certain change will lead to certain results.

An example may be helpful. In the previous section entitled **Conceptualizing**, we may assume that Bill was operating a drink stand selling orangeade and lemonade. Let us assume that after a week of operation he had the following information, or data.

1. Some days he lost money, other days just broke even; and some days he made a little.
2. Each day he sold more orangeade than lemonade.
3. At the close of each day, he threw out a lot of lemonade.
4. At the close of each day, very little orangeade was left.
5. Because of the cost of lemons and sugar, his costs for lemonade were greater than the cost of orangeade.
6. Many of those who bought lemonade seemed to prefer it only slightly over orangeade.

Examining the above information, it seemed to him that selling orangeade was profitable; selling lemonade, of which he threw away a lot each day, led to losses, and few customers really preferred lemonade much over orangeade. Bill now had the basis for his hypothesis, which he may have stated in this way:

> If selling orangeade is profitable; if selling lemonade is not; and if most customers like orangeade about as well as lemonade; then if I specialize by selling only orangeade, I will make a profit.

Bill had stated a **tenable** hypothesis because it was based upon, and was consistent with the information he had to work with. Bill had made a **prediction of a future outcome** conditioned upon an action he could take. Bill had stated a **testable** hypothesis, because he could specialize in orangeade, and see whether or not he made a profit.

Bill's next steps consistent with his hypothesis would be to specialize in orangeade, and to see whether or not after a period of time he had a profit. These actions are the test of his hypothesis. If he, in fact, had made a profit, he would say that his hypothesis was **confirmed**, and he probably would continue his speciality. If he, in fact, broke even or had a loss, he should **reject** his hypothesis, and start over again. Using the total information he now has, he should develop a new hypothesis and test it.

Examples

1. **Data Available:**
   a. A mother had three girls whose ages were 6, 4 and 2.
   b. When the oldest started to school, she bought several dresses, each made of very durable fabric, and each quite expensive.
   c. The eldest outgrew the dresses in one year.
   d. The second wore the dresses one year when she became 6.
   e. The third wore the dresses one year when she became 6.
   f. The first loved the dresses; the second complained that they didn't look like the other girls' dresses; and the third objected violently to their style.

Hypothesis

A. If the dresses made of durable but expensive fabric lasted over one year's wear by each of the three girls, and if the younger girls complained about their being out of style, then this mother could reduce the complaints by buying dresses of much less durable fabric at a much lower cost for each girl each year.

Is this next hypothesis also tenable and testable?

B. If the dresses made of durable but expensive fabric continued to be of good basic quality, but caused complaints by the younger girls because they were out of style, then this mother could reduce the complaints by remaking the dresses before they were worn by each successively younger girl.

Tenable: __X__ Yes _____ No

Testable: __X__ Yes _____ No

Note: If the purpose is to reduce the complaints of the younger girls, there may be other factors to be taken into account before the mother should decide which course of action to follow, that is, which hypothesis to test. Among such factors are the cost/quality of the cheaper dresses and the time/cost of remaking the expensive dresses. It is quite common in practice that such "other factors" enter into the decision of which of two or more equally tenable and testable hypotheses actually should be tested.

2. A real difficulty faced by students in developing tenable and testable hypotheses is determining which data among a mass of available data to use to develop a tenable and testable hypothesis. This example illustrates the problem.

Data Available:

From advertisements in the daily paper, the following information was obtained about stores selling lawnmowers:

Store	Brand	Size	Other Information
_____ 1.	A	21"	No service available
_____ 2.	B	18"	Self-propelled
_____ 3.	C	24"	Beautiful blue
_____ 4.	A	21"	Full service free for 1 year on mowers sold
_____ 5.	A	21"	Service available free first 6 months on mowers sold.

Hypothesis:

If a certain piece of goods has a value, and if service has a value, then the sales price of lawnmowers at the different stores will increase as the service offered increases.

Put an "X" in front of the stores whose advertisements, when taken together, make the above hypothesis tenable and testable.

Answer: 1, 4, 5 (All sell the same brand of mower in the same size, and differ among themselves only in terms of the amount of service offered without charge. Stores 2 and 3 sell other brands in other sizes and give no information about service. The hypothesis is testable by going to these same stores and obtaining their sales prices. If it is determined that the price at store 1 is least, at store 5 somewhat higher, and at store 4, the highest, the hypothesis will be confirmed.)

F: Generalizing

Generalizing is the process of developing a statement of relationships among a number of items or events, or the process of applying a known broad principle that can explain the occurrence of the items or events given.

Note: The measurement of this process requires that the item furnish a number of details, or that it refer to a set of details which the writer is practically certain that the students already possess.

Examples

1. A family bought a new household appliance. They paid in cash all but $100.00 of the total cost. They agreed to pay the $100.00 and carrying charges in payments of $20.00 per month. The store gave them this schedule:

Payment Number	Total Payment	Paid to Balance Due	Paid to Carrying Charge
1	$ 20.00	$ 18.50	$1.50
2	20.00	18.77	1.23
3	20.00	19.06	.94
4	20.00	19.34	.66
5	20.00	19.63	.37
6	4.77	4.70	.07
Totals	$104.77	$100.00	$4.77

In the space below, write one statement which describes the changes that take place in the share of each payment that goes to pay off the balance, and in the share that goes to pay carrying charges as this type of credit account is paid off.

Answer: _____

2. Malnourished children whose mothers prepared few regular meals ate mostly chips, bread and jelly, sandwiches made with cold cuts and drank bottled soft drinks.

   The children were brought into a school program providing breakfasts with fruit, cereal, or eggs, toast and milk and lunches of cooked meats, potatoes, vegetables, salads, bread and milk. They ate the bread, potatoes if french fried, threw the rest of the food in the waste can, and bought bottled drinks from the store across the street.

Which of the following generalizations best explains the children's actions?

A. Children will not eat foods they have not learned to like.
B. Mothers tell their children not to eat what they were given.
C. School lunches cost money.
D. The food would have made them sick.

PRACTICE EXERCISE 9

In this exercise, you are given a section of content entitled YOUR BREAKFAST CEREAL. Then you will be asked to write test items illustrating each of the six cognitive processes presented under Enabling Objective 8.

YOUR BREAKFAST CEREAL

What did you have for breakfast this morning? Was it cereal from a box you or your parents bought from a grocery store?

We are going to pretend that you did have cereal, and that it was purchased from a store. Long ago, you would have had fewer choices to make. There were fewer companies that produced cereal in boxes, and not as many stores selling it. The companies didn't make cereal with the sugar already in it. People had less money to spend on the purchase of cereal, too.

But this is today, not long ago. What happens when you decide you want cereal for breakfast? First, you must decide how much you wish to spend. Then, you need to know something about the cereal you are thinking of buying. How did you find out about it? Does anyone you know eat it? Where can you get it?

Once you think you will like it and want to buy it, you may have a few more questions about the purchase of a particular cereal. Some cereals contain more vitamins than others; some have sugar in them and others do not. Cereal boxes are different sizes. Some boxes have prizes in them. All of them cost money but each different box does not cost the same amount. Can you think of some reasons why cereals cost different amounts?

Using the above content:

A. Write one **recall** item

B. Write one **classification** item

C. Write one **inference** item

D. Write one **generalizing** item

E. Write one **conceptualizing** item

F. Write one **hypothesizing** item

Compare the items with one the author wrote. Do they seem to be similar in terms of process?

Illustrative Items

A. Recall

According to this story, what is the **first** thing you must decide when you start to choose which cereal to buy? Place an "X" in front of the right answer.

_____ A. which kind you will not buy
_____ B. whether or not your friends eat cereal
_____ C. how much you wish to spend
_____ D. how big a box to buy

B. Classification

Some of these things are **true now** and some **were true long ago.** Make an "X" in front of each thing that is **true now.**

_____ A. fewer choices of cereal
_____ B. more companies produce cereal in boxes
_____ C. fewer stores selling cereal in boxes
_____ D. sugar was not already in the cereal
_____ E. people have more money to spend on cereal

C. Inference

Although the story did not say exactly, I think the statement below that I have marked "X" is true.

_____ A. kids long ago ate more cereal than today
_____ B. it was harder to choose cereal long ago than it is today
_____ C. people eat more kinds of cereal today than they did long ago
_____ D. kids do not ask their mothers to buy cereal so they can get prizes

D. Conceptualizing (Concept: choice)

The big idea is that people must choose what they will buy if there are many things they could buy, because they don't have the money to buy, or don't want, everything. Put an "X" in front of the story below that describes a situation in which choosing is most necessary.

_____ A. Mr. Brown's store sells 15 kinds of cereal. He has most cereals in two or three sized boxes. Some have sugar, some vitamins, and some prizes.

_____ B. Mr. Jones' store sells rolled oats, corn flakes, wheat and rice cereals. All are in 16 ounce boxes and are just plain cereal.

E. Hypothesizing

If there were fewer cereals to choose from long ago than there are today, the older people ate the same kind of cereal much of the time when they were little children. Place an "X" in front of the statement below that identifies an appropriate hypothesis.

_____ A. There were fewer cereals long ago because the tastes of people were more limited.

_____ B. There were fewer cereals long ago because the incomes of people were less.

_____ C. Children today have sweeter tastes than they had long ago.

F. Generalizing

Make an "X" before the statement that is true.

_____ A. all companies making cereal make all kinds

_____ B. if you try, you can buy one cereal that has everything you want

_____ C. big boxes of one cereal always cost more than little boxes of any other cereal

_____ D. there are many more things to think about when you buy cereal today than there were long ago.

ENABLING ELEMENT 9

Norm Referenced and Criterion Referenced Tests

When a test has been prepared, administered to the students, and the papers scored, the teacher obtains a test score, and only a test score, for each student. Each test score is a bit of information to the effect that the student earning it answered that many of the questions on the test correctly, or received that many credits for his efforts. To make a meaningful decision about the attainment represented by the test score, the teacher must compare it with some generally understood and accepted standard. The process of comparing a test score to some stated standard commonly is known as referring the score to that standard.

In current practice, test scores are referred to one of two types of recognized standards. They may be referred to the mid-score, or to the median of the distribution of scores, earned on the same test by a group of students of the same general age and grade in school. The mid-score, or the median of the distribution of the scores of the group, is defined as the normal score. The group is called the norming group. When scores of individual students are compared with norm scores, it is said that they are being referred to a norm. Tests that provide a table that enables the teacher to compare the scores of individual students with norms are said to be "norm referenced" tests.

The second type of standard with which individual scores may be compared is a pre-set and stated number or percentage of items answered correctly. That number, or percentage, is called the criterion score. Tests which provide a pre-set and stated number of items that must be answered correctly for a student's performance to be accepted as satisfactory are said to be "criterion referenced" tests.

Students' scores obtained on a good test may be used in decision-making by referring them either to norms or to criteria. The same test may be either norm referenced or criterion referenced. For deciding which of the students are in greatest need of special help, or deciding whether or not help that has been given has been effective, a norm referenced interpretation should be made. For deciding who has, and who has not, attained a sufficient level of mastery to enable the student to go on to the next unit, a criterion referenced interpretation should be made. Thus the kinds of evaluative decisions to be made determine whether the teacher should interpret test results on a norm referenced or criterion referenced manner. The following example may be helpful.

Example

The students were studying a module on the computation of interest and other types of finance charges. This module normally would be followed by one requiring the students to compute alternative costs of purchasing a home and equipping it with certain high cost items if each of two financing plans were used.

The teacher would need to make two types of evaluative decisions. The first decision was which of the students had, and which had not, attained a sufficient

level of mastery to enable them to work with the content of the next module. The second decision was whether or not a special assistance program used by the three students who typically scored lowest in the class had been effective.

The teacher prepared a good ten-item test on the module being studied. He/she made the judgment that all students answering correctly seven of the ten items were prepared to undertake the next module. By making this judgment, the teacher set the criterion, and would be able to refer the score of each student to it. The test could be used in a criterion referenced manner.

The three students who had received the special assistance program typically scored lowest in their class. By making the rest of the class the norm group, that is, the group with whom their performance would be compared, the test could be used in a norm referenced manner.

After scoring the students' answers, the results were tallied as follows: a zero tally mark (0) represents one of the three special help students and a straight tally mark (/) represents one of the other students.

Test Score	Students	
10	//	
9	///	
8	~~////~~ ///	
7	~~////~~ /	0 Criterion
6	/	00
5	//	
4		
3		
2		
1		
Total	25	

In the example how many met the criterion? Yes, twenty. How many did not? They would have further work before attempting the next module. Do you see that using classroom tests in a criterion referenced manner assists in the making of important evaluative decisions?

In the example the three students who had received the special program are represented by "0's." Typically, their test scores were the lowest in the class. Their program could be judged to have been effective if they no longer scored lowest. Was the program effective? How does the use of the test results in a norm referenced manner show you this? Do you see that using test results in a norm referenced manner may help in the making of important evaluative decisions?

ENABLING ELEMENT 10

Reasons for Considering Other Observational Data

There are two basic reasons for considering observational data other than test scores when evaluating student attainment. First, the objectives for most units of instruction include some that can be assessed only by observing behavior other than test performance. Developing interests, changing attitudes, applications of what is learned to out-of-school situations, independent exploration of related topics, etc., are examples of such objectives. Behaviors related to such objectives must be observed when they occur and cannot be tested for formally.

Second, some behaviors seem to be indicative of change in cognitive learning, and many teachers feel that their occurrence should be weighed when evaluating student attainment. Participation in group or class discussion by withdrawn children seems to be indicative of improved cognitive learning. Bringing a picture or a short article pertaining to some phase of the topic and sharing it with peers may be a signal of better work to come. These behaviors too must be observed when they occur, and cannot be formally scheduled without losing their spontaneity and significance.

Behaviors which are the subjects of stated objectives should be observed and fully considered when evaluating student attainment. Behaviors thought to be indicators of change for better or for worse in the level of student attainment likewise should be observed and considered when devising teaching strategies and when evaluating student attainment.

ENABLING ELEMENT 11

Techniques for Recording Other Observational Data

Observations of unscheduled student behavior typically should be recorded as anecdotes for use primarily by the teacher making the observations. For evaluative uses, they need be preserved only for relatively short periods.

As anecdotes, they should report as precisely as possible just what the student did, that is, behavior. A brief notation recording the setting, or circumstances in which the behavior occurred usually is helpful. Involved inferences about the students' motives, or descriptions of the impact upon other students, the class or the teacher should be omitted. In terms of style, the entries should be informal, but clear.

The records may be kept in a simple notebook, or in an ordinary tablet. It is good practice to record on the first page the types of behaviors to be looked for because they are directly stated or implied in the unit objectives. Separate from these, but clearly noted, should be those behaviors which the teacher believes are indicators of significant changes in student learning, and hence worthy of observation. If the combined list is brief, it is more probable that the teacher actually will be able to keep up with the task and end up with good data to work with.

It is recommended that the students' names be entered alphabetically, one name to a page. The ruling of each page should provide one column for entering the date of each behavior recorded, and the remainder of the page probably should be blank. Attempts to record anecdotes on elaborately ruled pages rarely are successful and often are confusing. Each teacher may experiment with a recording system until a workable one is developed.

POST-ASSESSMENT

1. A teacher wrote an item to measure a certain complex cognitive skill. 60% of the students marked the correct answer. The results established that:

   A. 40% of the students tried to use some other skill
   B. 60% of the students got the answer correct by some process
   C. 40% of the students could perform the skill but lacked basic knowledge
   D. all who marked it correctly can perform the skill.

For each of the next two items, select the letter from the list below of the cognitive skill or process that students probably would use in responding to it. You may use any skill or process as often as you wish.

Skill or Process

A. generalizing
B. inference
C. classification
D. hypothesis development

2. Given the following facts:

   a. insurance companies pay 3.5% interest to policyholders,
   b. national banks pay 4.5% on savings,
   c. savings and loan associations pay 5.0% on savings,

   If a person needed to borrow $10,000.00 to purchase a $50,000.00 home, which type of lending agency probably would lend him the money at the lowest interest rate?

Skill or process numbered _____ probably would be used by students responding to this item.

3. A girl found that she could invest in each of the following types of securities at the interest rates shown below:

   U.S. savings bond–3.75%
   national bank savings account–4.5%
   savings and loan account–5.0%
   first mortgage on a home–6.5%
   second mortgage on a home–8.75%

In one short sentence, what probably accounts for the increase in interest rate as the types of investment change as you read down through the list?

Skill or process numbered _____ probably would be used by students responding to this item.

4. A teacher told a student that his reading comprehension score indicated that he was reading two grades higher than his grade level.

Obviously, the teacher was interpreting the student's test score in:

A. an unethical manner
B. a criterion referenced manner
C. a norm referenced manner
D. an erroneous manner

5. Which two of these possible reasons really are NOT proper justifications for considering observational data other than test scores when evaluating student attainment?

NOT Proper Justifications

A. 1, 2

B. 2, 3

C. 3, 4

D. 2, 4

Possible Reasons

1. Certain objectives can be assessed only in terms of the behaviors of students in the course of their activities.
2. Some students do not do well on tests.
3. Some things students do probably indicate that changes in their learning are taking place before the results show up on test scores.
4. Observations of other behaviors give the teacher a defensible basis for awarding better grades to some students and lower grades to other students.

6. The following paragraph contains statements describing a technique for recording observational data other than test scores. Not all of the statements may be correct. Read the paragraph carefully and choose the alternative which lists the numbers of the statements which are NOT correct.

The Numbers of the Statements that are NOT correct

A. 3, 5

B. 2, 6

C. 1, 4

D. 2, 4

A Technique for Recording Observational Data

1. Observations of unscheduled student behavior should be recorded as anecdotes, primarily for use by the teacher making the observation.

2. Each record should state just what the student did and contain a brief notation of the circumstances under which the behavior occurred.

3. Each entry also should contain the teacher's inferences about the students' motives and the affective effects upon other students.

4. Prior to making the actual observations, the teacher should decide just what behaviors she will look for, and list them, keeping the list brief.

5. The behaviors to be observed may include any that interest the teacher, because she has no way of knowing to what she will relate them.

6. A simple notebook page for each pupil, listing his name, providing a column for dates, and blank space for recording the observed behaviors probably is the best record system to start with.

Answers: 1-B; 2-B; 3-A; 4-C; 5-D; 6-A

ENABLING ELEMENT 12

Ways of Using Observational Data

There is no single way of using observational data in the making of evaluative decisions. There are no ways that are universally correct, and probably none that are universally wrong. The ways to use them are matters for the judgment of each teacher, and they should be arrived at judicially, taking into account the situational factors in each case and the quality of the data available.

Two types of mistakes may be made in dealing with observational data. The first is to try to assign numerical values to the observations and then to treat the values assigned as if they were measures of physical objects. They are not, and treating them as if they were is a serious mistake. Even a simple count of the number of observations of a given behavior may be misleading because the number of times it occurred and was not observed cannot be known. The second type of mistake is to attempt a completely unplanned approach to the use of observational data when evaluating student achievement. The results will be difficult to explain, and the teacher may be accused of being too severe with some students and too generous with others. Some reasonable rationale for the use of observational data in assigning student grades and in evaluating learning/teaching strategies is a practical necessity.

In many situations, such observational data should supplement test data and grades for class work. In effect, they may furnish the basis for adding a few points or for subtracting a few points. Usually either of these can be done with good judgment. Another very important use is in explaining to a student or to parents in a counseling setting why the student is achieving as he is, and identifying things he could do to improve. Used in these ways observational data should make positive contributions to evaluation and to learning.

It usually is advisable for the teacher to begin to use such data by considering how best they may be used with the task at hand. From this deliberation, the teacher should develop a few guidelines. By referring to the guidelines from time to time, the teacher will maintain an even-handed usage throughout the task. With experience in formulating personal guidelines and in applying them, the real value of such observations almost certainly will increase.

However, there are some behaviors which are highly valued in consumers' education, but simply are not observable—life-long personal budgeting, perhaps. At any given time and in any given setting, observational data will not be complete, but it should be better than none!

PRACTICE EXERCISE 10

Other Observational Data

Place a check before each of the statements below which is correct.

_____ 1. The assessment of many important objectives requires observations of behaviors other than test scores.

_____ 2. Some behaviors are clues to important changes which students may be making.

_____ 3. Observations of what students actually do should be considered when assigning grades.

_____ 4. It is not necessary to keep a record of such behaviors because the teacher will remember them, anyway.

_____ 5. A short list of behaviors to be observed helps the teacher by reminding him/her of the important things to be looked for.

_____ 6. Records of observations should be brief anecdotes.

_____ 7. Each teacher should experiment with ways of keeping records of observational data until he/she develops one that works well.

_____ 8. Records of observations of students' classroom behaviors should be given numerical values and treated as test scores.

_____ 9. Records of observations of student behavior are valuable for student and parent conferences.

_____ 10. Planning how he/she will use observational data will improve the use a teacher actually makes of them.

Key for Practice Exercise 10: 1, 2, 3, 5, 6, 7, 9, 10.

Organizational Note

Up to this point, the contents of this module have dealt with the obtaining of data pertaining to student attainment. The material was directed to the evaluation of the achievement of the individual student.

A professional responsibility of teachers is that of evaluating, making decisions about, the effectiveness of instructional sequences, units or modules as they actually were carried out in the classroom. The purpose of these evaluations is to enable the teacher to make changes in the instructional strategies, in the instructional materials provided to the students, and in the teacher's own activities as the study and instruction progresses, or before the same topic is taught again.

Meeting this professional responsibility requires an evaluation of three components. The first is the students' activities—the types of study in which they engage, time spent in each, estimates of the effectiveness of each for students of different types, and contributions made by instruction in other classes and by out-of-school experiences.

The second component is the instructional materials provided. The evaluation should take into account their adequacy, availability, and suitability for the students involved.

The third component is the input of the individual teacher. The teacher's input largely is determined by a basic grasp of the subject matter, by planning efforts, by activities during the instruction, and by the demands made upon the teacher for related and unrelated work with persons and groups other than the students.

Within this module, the evaluation of the students' activities and of the instructional materials will not be considered because of time limitations. The module will close with a very brief treatment of the evaluation of the input of the teacher.

ENABLING ELEMENT 13

Reasons for Documenting Teacher Input

It is almost universally accepted that what the teacher puts into a learning/teaching sequence influences directly what the students receive from their study. But the teacher may put in many things. Not all of them are obvious to students, and some may not be recognized as inputs by the teacher. The first reason for documenting the things a teacher does is to make them evident to the teacher, and to others who have reasons to be interested and concerned.

It is a fact of life for teachers that they have only so much time and energy that can be devoted to any given teaching task. The effectiveness of their instruction is dependent on a very considerable extent upon the way they distribute their time and effort among the principal activities in which they engage. The second reason for documenting the things teachers do is to enable them to readjust the distribution of their time and effort so that their total input may be more effective.

In the final analysis, the basic reason for a teacher to record the types, the amounts of time, and possibly his/her judgments about the effectiveness of the activities related to a module, learning sequence, or unit is to enable the teacher to teach students more effectively.

ENABLING ELEMENT 14

Illustrative Form for Documenting Teacher Input

The form in which teacher input should be documented is that which the teacher concerned finds usable and useful. Therefore the form presented below is illustrative of the types of information which are useful in making evaluative decisions, and simple enough to be usable in most situations. Teachers should experiment by modifying the form until they develop one which they find most useful and usable.

The form presented lists three broad classes of teacher inputs, namely: A. Planning and Direct Preparation; B. Direct Instruction; and C. Other Related Activities. Each of the broad classes is sub-divided into more specific activities related to it. For each of the specific activities in which the teacher engaged, space is provided for recording the dates and the time spent on each date on that activity. The last column provides space for recording the teacher's notes and comments about the conduct of each activity.

The documented input, if kept current, may be very valuable for formative evaluation purposes. For example, by glancing over the previous day's entries the teacher can identify activities on which too much time was spent and those on which too little time was spent. As a basis for the summative evaluation of the learning/teaching sequence, the teacher may add the number of minutes spent on each activity and examine the grand total and the balance among activities. Relationships between the kinds of things the students produced and their levels of achievement and the types and extent of teacher input probably will become apparent.

Documenting teacher input serves its purpose when it enables the teacher to adjust teaching activities while instruction is in progress, or to adjust the distribution of teacher efforts the next time the unit is taught.

Figure 3. Teacher Input Form

Consumers' Education

Name Topic Class

A. Planning and Direct Preparation

Date	Time Spent	Activity
		1. Planning teaching strategies and student activities.
		2. Reviewing Content
		3. Assembling Materials
		Total time spent in Planning and Direct Preparation (add column 2)

B. Direct Instruction
Consumers' Education

Date	Time Spent	Activity
		1. Large group instruction
		2. Small group instruction
		3. Individual instruction
		4. Evaluation (test preparation, scoring, grading, observing students at work, etc.)
		Total time spent in Direct Instruction (add column 2)

C. Other Related Activities
Consumers' Education

Date	Time Spent	Activity
		1. Formal inservice classes and seminars
		2. Independent study of the topic taught
		3. Related committee work
		4. Related community work
		Total time spent in other related activities (add column 2)

PRACTICE EXERCISE 11

1. Which of the following are valid reasons for a teacher to make a record of his/her activities? Check as many as are correct.

   ____ A. to help the teacher distribute his/her time wisely
   ____ B. to prove that work was done
   ____ C. to call to the teacher's attention the types of related activities in which he/she did or possibly should have engaged.

2. Classify each of the entries on the right by recording its number beside the type of teacher activity to which it belongs:

Type of Activity

A. Planning and Direct Preparation
   ____ a. Planning teaching strategies and student activities
   ____ b. Reviewing content
   ____ c. Assembling materials

B. Direct Instruction
   ____ d. Large group
   ____ e. Small group
   ____ f. Individual
   ____ g. Evaluation

C. Other Related Activities
   ____ h. Formal Inservice
   ____ i. Independent study
   ____ j. Related committee work
   ____ k. Related community work

1. worked with Susy
2. made plans for teaching "banking"
3. scanned chapter in the textbook
4. prepared program on Consumers' Education for parents
5. attended extension class
6. prepared test for unit
7. worked in library selecting materials for the class

Check your answers with the key below.

Key to Practice Exercise 11:

Item 1. X A. ___ B. X C.

Item 2. 2 a. 3 b. 7 c. ___ d. ___ e. 1 f. 6 g. 5 h. ___ i. ___ j. 4 k.

POST-ASSESSMENT

1. Listed on the right are possible ways of using observational data. Choose from the alternatives on the left the one that lists the ways of using observational data that are NOT generally advisable.

NOT generally advisable ways

A. 2, 3

B. 1, 5

C. 3, 4

D. 2, 3, 4

Possible Ways

1. Quantify them and treat the results mathematically. 2. Plan how best to use the records for all students if they will be used in arriving at the grade for any student. 3. Use the records for counseling individual students and in conference with individual student's parents. 4. Use them as one part of a student's record, being aware that the behaviors recorded constitute only a part of the total of any student's behaviors. 5. Use them in a completely unstructured and unplanned manner so you will not prejudice the behaviors of any student.

2. Which of the Possible Reasons listed is (are) NOT valid professionally as a reason(s) for documenting teacher input into the learning/teaching process?

NOT valid

A. 1

B. 1, 3

C. 3

D. 2

1. Documenting the several things the teacher does makes them evident to the teacher.
2. Documenting the several things the teacher does provides evidence that the teacher did do something.
3. Documenting the several things a teacher does enables him/her to redistribute his/her efforts and thus become more effective.

Classify the specific teacher activity given in each of the two following questions by recording the letter number of the type of teacher input to which it belongs from the list on the right.

Specific Teacher Activity	Type of Teacher Input
3. Prepared presentation on Consumers' Education for parents' club (2 hours)	A. Related Activities: Committee B. Related Activities: Community C. Planning: Assemble Instructional Materials D. Direct Instruction: Large Group
4. Scanned text materials for the unit (30 minutes)	A. Planning: Assemble Materials B. Other Activities: Formal Inservice C. Planning: Review of Content D. Direct Instruction: Individual

Answers: 1-B; 2-D; 3-B; 4-C

GROUP ACTIVITY

1. Have teachers meet in small groups and analyze some tests that have examples of poor test items. Give reasons for their statements.

2. Have teachers meet in small groups and construct well constructed test items. Allow time for large group discussion.