DOCUMENT RESUME

ED 129 348 JC 760 509

TITLE	Handbook for Evaluating Objectives.
INSTITUTION	Saint Louis Community Coll., Mo.
PUB DATE	Aug 76
NOTE	15p.
EDRS PRICE	MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS	*Attitudes; Behavior; Departments; Educational Objectives; Evaluation; *Evaluation Methods; *Guides; Manuals; Measurement; *Objectives; *Program Evaluation

ABSTRACT

This document was compiled for the Participative Management Program at St. Louis Community College. It is composed of four parts, each dealing with an aspect of evaluation of objectives. Part 1, An Introduction to Evaluation, discusses the difference between measurement and indication, the three general evaluation types (internal data, surveys, tests), their dependability, and the conceptual steps, tactics, and application of the three types. Part 2, Program Evaluation: Quantitative Methods, discusses human and theoretical problems with quantitative data, defines levels of measurement, and illustrates sample objectives measured at different levels. Part 3, Introduction to Attitudes and Attitude Measurement, defines attitudes and their cognitive, affective, and behavioral components, discusses functions of attitudes and a rationale for their measurement, describes characteristics of attitudes and behavior, and provides guidelines for measuring attitudes through use of examples. A general discussion of attitude measurement and the expected value to educators is included in this part. Part 4, Evaluation of Support Department Objectives, discusses objective formulation, emphasizing the writing of measurable objectives that are realistic, attainable, and focused on results. It is suggested that the writer of objectives decide on a quantitative standard, define the audience, set a deadline, and determine the mechanics of measurement for each objective. (JDS)

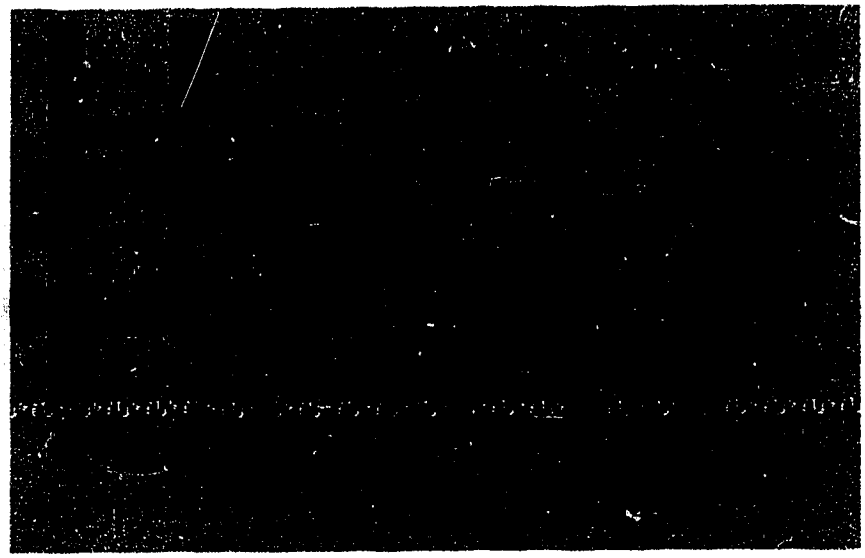

TABLE OF CONTENTS

Participative Management Program
St. Louis Community College
August, 1976

PART 1
An Introduction to Evaluation

Lawrence Klein

This introduction to evaluation is intended to ***suggest*** possible approaches to evaluation that each group writing objectives may want to consider. Some objectives may seem so important that several different evaluation methods will be chosen, while at the other extreme, some objectives may seem so secondary or tangential to a group's purpose as to not warrant evaluation at all. Moreover, some objectives may be deemed worthy of evaluation every semester, while others need evaluation only every few years. Finally, an appropriate evaluation technique may be determined but the technique may be too costly (money, people or time) to pursue.

I. Measurement * or Indication **?

If the method of evaluation chosen is expected to produce an exact statement of accomplishment for an objective or implementation event, then it should be listed as ***measurement***, which would mean it is expected to be an accurate way of measuring the objective.

But if a method does not measure the accomplishment of all of an objective, or if the method is not felt to be completely accurate, it should be listed as an ***indicator***. This would mean that it describes something about the accomplishment of an objective, but not all that needs to be known.

II. The three general evaluation types are:

1. Internal data - data that is already collected or whose collection can be set up as a routine clerical function. Examples:

   grade reports
   enrollment figures
   attendance counts - for libraries, laboratories, offices, social and athletic events
   cash register receipts
   number of telephone calls
   number of complaints, letters of commendation
   number of donations, amounts of donations

2. Surveys - responses from graduates, former students, present students, prospective students, employees, advisory board members, and local residents can be solicited through surveys conducted either regularly or occasionally.

3. Tests - mental and physical skills, memorized information, and task performance can be evaluated by asking students, employees, etc. to answer questions or do a task.

* Measurement - used for those objectives for which precise quantitative and qualitative standards can be determined.

** Indicators - used for those objectives for which some data can be identified which provide indications of achievement.

III. How dependable are any of the above types of evaluation?

Their dependability is a function of how accurately data is collected, how appropriately it is analyzed and how carefully it is interpreted. The number of people who call the college president to complain or compliment may be an important number, but it can't be interpreted as an accurate evaluation of attitudes throughout the community.

Similarly, survey responses may not be entirely honest, may not come from a sufficiently large and representative group, or may not reflect an adequate understanding of the questions asked.

The validity and reliability of particular tests must be established here within the college. Tests raise the additional problems of nervousness, illness, and poor test-taking skills.

Considering these drawbacks and other potential hazards, the first rule of evaluation, ***no matter what method is used,*** is that results, even if expressed in exact numbers, are ***approximations.*** All too often, however, this is forgotten, because the precision of numbers is intellectually seductive.

One way of avoiding the trap of giving too much credence to a particular set of numbers is to use a ***combination*** of evaluation methods. Two or more different tests or surveys, ***combined*** with some internal data will possibly

1. cause people to pause before jumping to a conclusion based on one set of seemingly precise results
2. offset the chance of a poorly conceived or constructed evaluation method having too much effect.

The problem with such combinations is that they are confusing to work with, but perhaps our coping with such confusion will lead us to methods of evaluation that we trust as generally accurate. The alternative is to search fruitlessly for simple answers to complex questions.

IV. Conceptual Steps toward evaluation:

1. The first step is defining the objective, the hoped-for end result. This process is already well underway.
2. Next, the steps that lead to the objective need to be defined. The writing of "implementation events" can be seen as such a definition.
3. Throughout the process, definition of implementation events needs to be reviewed, since an implementation event, fully accomplished, may be found to be ineffective in accomplishing the objective. Such review can be likened to the formal process known as "successive approximation".
4. Finally, the starting point or present level of accomplishment needs to be determined. How many people begin a program, at what "entry level"?

V. Tactics for the three general evaluation types:

1. Internal data (general information collected on a regular basis) needs to be presented clearly. One way of doing this is constructing graphs or other ***visual*** displays that may reveal more than the numbers alone would.
2. Surveys are requests for information and attitudes. The considerations involved in constructing and administering a survey that will be both valid and reliable are highly complex.

   For attitudinal measurement, three techniques are generally accepted: Likert Scales, Osgood's Semantic Differential, and Thurstone Scales. These techniques are explained on pp. 17-21 of the ***Guide for Evaluation*** on reserve at each library.

   For further help on survey construction and administration, consult with the following deans:

Pete Hirsch	Florissant Valley
Ron Lingle	Meramec
Lana Weinbach	Forest Park

3. Tests can be written, oral, or task performance (or a combination).
   a. Written tests, whether "objective" or "open-ended" demand that the test-taker respond appropriately to a question or command, so the first requirement is that questions and commands be clear and unambiguous. Test-takers should be led to believe that they can perform well on the test; otherwise, their fears of failure may obstruct their doing their best.
   b. Oral and performance test questions or commands also need to be clear, but the test-taker has an opportunity to request clarification when needed.
   c. Performance tests (typing, shorthand, athletics, drama, laboratory techniques, etc.) involve a problem that can best be described as "stage-fright". One way of reducing this is to give several opportunities for the same test.

   This discussion of tests is purposely superficial. Until more is known about test-taking, caution must be the by-word in test-making.

In addition to the deans mentioned above, the following people may be able to help with test construction

Phil Carlock	Forest Park
Bob Richey	Forest Park
Betty Duvall	Florissant Valley
Jim Pierce	Meramec

VI. Specific Applications

It is difficult to choose a college or program objective and suggest how it might be evaluated without seeming to be dictating to the college or program.

For suggestions that you are entirely free to reject, see the ***Guide for Evaluation,***

pp. 29-42 program objectives
pp. 43-53 and
Appendix A college objectives

For more suggestions, it might be productive to invite a small group of people from other programs within the college or from other colleges to give their ideas on evaluating a specific objective. These quest interviews might tell a program or college writing team how their objectives are being read by others, and how others view their role in the total effort of St. Louis Community College.

PART 2
Program Evaluation: Quantitative Methods

James O'Grady

INTRODUCTION

Part I of this handbook noted that some program objectives may seem so important that several different evaluation methods will be chosen, while at the other extreme, some program objectives may seem so secondary or tangential to a group's purpose as to not warrant evaluation at all. For those departments who want to evaluate program objectives quantitatively, the following information covering problems, levels of measurement, and specific examples, may be of help.

I. Problems with Quantitative Data

Human

1. Bias - Persistent error in one direction during the collection process.
2. Omission of important factor(s) - Omitting significant information.
3. Carelessness - Error introduced through improper collection, analysis and interpretation.
4. Non-sequitur - An inference that does not follow from the premises.
5. Non-comparable data - Comparing unlike measures.
6. Confusion of association with causation.
7. Poorly designed experiment.

Theoretical

1. Insufficient data - Adequacy of sample number.
2. Unrepresentative data - Adequacy of sample relative to population.
3. Concealed classification - Sub-groups in population weighting population.
4. Misleading totals - Combining unlike values.

II. Levels of Measurement

When data are being collected, the process of assigning a value or score to the observed phenomenon constitutes the process of measurement. The rules defining the assignment of an appropriate value determine the corresponding level of measurement. The different levels are distinguished on the basis of the ordering and distance properties inherent in the measurement rules. Knowledge of these rules and their implications is important to the evaluator because various statistical techniques are appropriate for data measured only at certain levels.

The traditional classification of levels of measurement was developed by S.S. Stevens. He identified four levels: nominal, ordinal, interval and ratio. A simpler classification than Stevens's is to divide variables into quantitative and qualitative types. Quantitative variables are those for which a fixed unit of measurement is defined--essentially variables at the interval and ratio levels. These are the variables for which the most powerful and sophisticated statistical techniques have been developed. Qualitative variables, then, are all others--namely, those at the nominal or ordinal level.

Another distinction, which can be made based on levels of measurement, is between parametric (or quantitative) and nonparametric (or qualitative) statistics. Nonparametric statistical procedures require few assumptions about the distribution or level of measurement of the variables and may be applied to nominal and ordinal data. The parametric procedures require more stringent assumptions concerning the distribution of the data and are designed primarily for data at an interval or ratio level of measurement.

Stevens's four levels of measurement and examples are presented as follows:

1. *Nominal or classificatory class* - Measurement at its weakest level, where numbers or other symbols are used to identify groups to which various objects belong. The scaling operation is partitioning a given class into a set of mutually exclusive subclasses with no high or low point. For example, we place graduates into distinct categories by program, Art, Business, etc. and do not imply that one is "greater than" or "better than" the other.

2. ***Ordinal or ranking scale*** - At the nominal level, objects in one category of a scale are shown to be different from the objects in other categories of that scale. At the ordinal level, the objects are different and they stand in some kind of relation to them. Typical relations among classes are: higher, more preferred, more difficult, more disturbed, more mature, etc. The ordinal level does have a higher and lower point but lacks equal intervals. For example, we may classify families of graduates according to socioeconomic status: upper, upper middle, lower-middle, and lower. We know only that one is greater or lesser than another but cannot say how much greater or lesser.

3. ***Interval Scale*** - At this higher level of measurement the characteristics of higher and lower points are present and in addition the distances between any two numbers on the scale are of known size. However, the zero point and the unit of measurement are arbitrary. For example, an academic grade point can be considered measurement on the interval level.

4. ***Ratio Scale*** - At this highest level of measurement, all the characteristics of an interval scale are present and in addition it has a true zero point as its origin. Normally, per cent measures are on the ratio scale.

III. Sample Objectives Measured at Differing Levels

Program Objective:	To reduce the attrition rate in the Art Program.	To improve student appreciation of music.	To serve persons requesting counseling services in a courteous manner.
Levels of Measurement			
Nominal	Classify students by who completed and did not complete program.	Classify students by indication of appreciation or lacking an appreciation of music.	Surveys of students will rate counseling staff favorably on this objective.
Ordinal	Classify and categorize students not completing program by possible variables, i.e., outside employment. Less significant, significant, and more significant.	Classify and categorize students who indicate an appreciation of music by possible criteria i.e., knowledge of composers. Completely unfamiliar, familiar, very familiar.	Surveys of students will rate counseling staff above average better on this objective. Inadequate, adequate.
Interval	Evaluate and Categorize students not completing program by academic grade point.	Evaluate students who are unappreciative of music by a score on the MAT.	
Ratio	Evaluate students not completing program by student credit hours of enrollment.		

PART 3
Introduction to Attitudes and Attitude Measurement

John Mukavetz

I. Some Definitions of Attitude

A. An attitude is a mental and neural state of readiness, organized through experience, influencing an individual's response to all objects and situations with which it is related.

B. An attitude is an idea charged with emotion which predisposes a class of actions to a particular class of social situations. Attitudes have three components.

1. ***Cognitive*** component - an idea or concept
   example: St. Louis Community College
2. ***Affective*** (emotional) component - an evaluative (good-bad) label
   example: St. Louis Community College is good and valuable to the community.
3. ***Behavioral*** component - a predisposition to action (seeking-avoiding)
   example: St. Louis Community College is good and valuable. Therefore, I would ***probably*** support legislation which would provide more resources for the College.

Measuring attitudes involves measuring statements about beliefs or categorizations (cognitive component), statements about emotions and feelings (affective component), and ***statements*** about actions (behavioral component). The predicted behavior is not guaranteed.

II. Functions of Attitudes (Why measure attitudes?)

A. Attitudes help people to understand the world around them, by organizing, categorizing, and simplifying the complex and sometimes confusing input from the environment. Attitudes provide an overlay of PREDICTABILITY to our world.

B. Attitudes help people to protect their self-esteem, by making it possible to avoid unpleasant truths about themselves.

C. Attitudes help us to maximize rewards from the environment, since we get along best with people who have attitudes similar to the ones we hold.

D. Attitudes help people to express their fundamental values (Where do you stand on the abortion/right-to-life issue?).

III. Characteristics of Attitude and Behavior

A. Attitudes and behavior interact and are interdependent.

Attitudes and behavior affect each other and are affected by each other. For example, a student who has negative attitudes toward the study of philosophy may avoid lectures, books, and courses which deal with the subject. If the student finds that he has to take a philosophy course to fulfill a degree requirement, he may, through first-hand experience with the subject, change his attitudes toward philosophy. He may sign up for another philosophy class, do some outside reading, approach his other subjects from a different perspective. The feedback loop continues, as the student's attitudes change even more, to the point that he recommends philosophy classes to his friends.

In an alternative example, the presentation of the philosophy class may convince the student that his original attitude was correct (the attitude is reinforced), and no verbal or behavioral change will result.

B. Attitudes and behavior are situation-specific.
In a classic example, hard-rock coal miners in West Virginia demonstrated strong racial prejudices above ground. Within the mines, though, discriminatory verbal statements and overt behavior did not exist. The dangers of working in the deep mines and the necessity for cooperation and trust produced a temporary situation-specific attitude change.

C. Attitudes are learned.

Attitudes can be modified and changed according to the principles of learning theory. Teaching involved changing attitudes. For example, in a secretarial program, a student may show increased performance in typing and other business classes over a number of semesters. If, however, the student's attitude toward a secretarial occupation becomes more and more negative because of the impersonal and mechanical way the courses are taught, he or she may change majors and pursue a different occupation. This would be a waste of time and money for both the student and the department involved. Alternatively, the student may go into secretarial work, and because of intense negative attitudes, do a poor job.

D. Attitudes predispose behavior along a continuum, from

Giving simple statements of agreement or recognition → Giving statements of support → Giving statements about future actions → Proselytizing

IV. Measuring Attitudes

A. Some quantities to consider in attitude measurement

1. Magnitude - How much?
2. Frequency - How many? How often?
3. Latency - How long would it take before some action occurs?
4. Extinction - How long will the action continue, unreinforced?

B. There are many techniques and instruments available for use in measuring attitudinal objectives. Rating scales, questionnaires, check lists, and attitude - loaded objective tests may be used. See the attached (Appendix I) ***Informal Criteria for Attitude Statements*** (Girod, 1973) for some guidelines in writing attitudinal items. In general, first select major constructs, ideas, or concepts, then write as many items as possible related to those constructs.

Example Program - ***Management and Supervisory Development***

Attitudinal Objective	**Description of Student Behaviors**	**Methods of Evaluation and Measurement**
1. To improve attitudes toward women and minorities	**1a.** Statements of willingness to work with, promote, give raises to, give increased responsibilities to women and minorities **b.** Statements of willingness to explore the problems women and minorities have historically had in industry.	**1a.** Before and after measurements using Likert-type items (always/never, strongly agree/disagree). **b.** Use of Semantic Differential Techniques to measure changes in Ss concept structure. **2.** Before and after use of a rating scale requiring Ss to rank a series of activities according to Ss willingness to participate in them.

C. The problem of socially acceptable responses on psychological tests is well recognized. Various methods have been devised to control for this problem in attitude research. Some are discussed here.

1. The maintenance of honest and open communication between students and teachers will decrease the tendency for students to give the answers they think the teacher wants.
2. Make a point of requesting honest, first impression responses.
3. Limit the amount of time students spend on the evaluation instruments.
4. If at all possible, maintain the students' anonymity. Insure their rights to privacy. Attitudes are intimately held, and will be more honestly divulged in non-threatening situations.
5. Alternatively, tell the students exactly what you are measuring and why, tell them that one of your objectives is to change attitudes, and involve the students ***directly*** in the evaluation process. If done sensitively, involvement in the attitude measurement procedure may (probably will) in itself serve to change attitudes.

D. An example of an attitude measurement study using the Semantic Differential Technique (hypothetical)

1. According to factor analytic studies originated by Osgood (1957), essentially three major independent dimensions underlie the meaning of all ideas, objects, concepts, etc. These are:
   a. Evaluative dimension - An idea is good or bad, clean or dirty, fair or unfair, honest or dishonest, beautiful or ugly, believable or unbelievable.
   b. Potency dimension - An idea is strong or weak, big or small, powerful or powerless, heavy or light.
   c. Activity dimension - An idea is active or passive, hot or cold, fast or slow, alive or dead.
2. The connotative meaning of an object, idea, or concept can be mapped along the three dimensions by using a series of nine or twelve bipolar adjective scales (good-bad, strong-weak, fast-slow).
3. Evaluative dimension bipolar adjective scales are used to measure attitudes. Since this dimension accounts for over half of the variability in a factor analysis of the meaning of concepts, reliable and valid conclusions can be based on the results of the evaluative dimension scores alone.
4. Osgood (1957) reports test-retest reliability coefficients of .85 for meaning in general and .91 for attitudes. In comparison with other attitude scales, Thurstone, Guttman, and Bogardus Social Distance Scales, validity coefficients of .74 to .82, .78, and .72 to .80, respectively, were found (Smith, 1963).
5. A typical Semantic Differential, with instructions, may be found in Appendix II. A more thorough discussion of the Semantic Differential Technique is found in Snider and Osgood (1969).
6. Hypothetical Example - The present investigation was conducted to determine the feasibility of using Osgood's (1957) Semantic Differential (SD) in the evaluation of proposed institutional goals and objectives for staff being developed at Tinpan Alley Community College. Classified staff and administrators' attitudes toward the following 12 college objectives were compared:
   1) helping them develop a capacity for critical thinking
   2) helping them develop the ability to make independent judgments
   3) helping them improve their capacity for planning
   4) helping them improve their creativity
   5) helping them improve their capability for communication
   6) helping them with their career development
   7) helping them develop an improved capacity for earning a living
   8) helping them improve their appreciation for points of view different from their own

9) helping them develop improved abilities for relating to other individuals
10) encouraging involvement in further educational opportunities
11) realizing equal consideration
12) promoting participative decision making

Three measures generated by the SD scores were used to determine: (1) the general attitude of the two groups toward the objectives; (2) the relative "value" and "credibility" (believability, probable success), of the objectives within each group, and (3) if a "credibility gap" exists, whether there was a significant difference between the two groups.

RESULTS

Part 1 - Both classified staff and administrators judged the proposed objectives favorably. For classified staff, attitudes were significantly favorable (different from neutrality) for all objectives except No. 11 (realizing equal consideration). For administrators, attitudes were significantly favorable for all objectives except No. 4 (improve creativity), No. 7 (improve capacity to earn a living), No. 11 (realizing equal consideration), and No. 12 (promoting participative decision-making). Overall, administrators judged all objectives less favorably than classified staff. When the objectives were ranked according to their relative importance for the two groups, large differences were found. The rank correlation of .483 was not significant.

Part 2 - For all objectives, classified staff's "credibility" scores were less favorable than "value" scores, and significant differences were found for five objectives. Administrators' "credibility" scores were also less favorable than "value" scores, and significant differences were found for six objectives. Both classified staff and administrators judged three objectives significantly less credible than valuable.

2. making independent judgements;
4. improving creativity;
9. improving ability for relating to other individuals.

Part 3 - Overall, the administrators' "credibility gap" score was higher than the classified staff for objectives No. 2, 3, 4, 7, 8, 10, 11, and 12. Significant differences were found for objective No. 2 (developing the capacity for making independent judgments) and three others approached significance. These appeared to be more job-skill related.

DISCUSSION

Generally, the results show that those persons responsible for developing college-wide objectives at a community college face an unenviable task. First, administrators, who may eventually be charged with implementation of the objectives, have consistently less favorable attitudes toward them than do the classified staff. Second, the relative importance of the objectives is greatly different for the two groups. A third indication of possible trouble ahead are the "credibility gap" scores, which indicate a cynical, "it may be good, but it will never work here" attitude. More prevalent among administrators, the scores may point to long-standing behavior patterns which may be difficult to change.

In summary, the results support the continued use of the Semantic Differential technique in the measurement of attitudes toward institutional objectives. Also pointed out were potential problems caused by the conflicting positions of staff groups in the overall evaluation and relative importance of goals and objectives, as well as the existing "credibility gaps" associated with the objectives. The consistency of the results strongly emphasizes the discrimination and sensitivity of the SD technique in the measurement of attitudes.

EXPECTED VALUE TO EDUCATORS

Given the current emphasis on evaluation and the movement in many areas of higher education toward a management-by-objectives approach, it is critically important to be able to evaluate staff, student, and community attitudes toward proposed objectives. Without such an evaluation, the institution may move blindly forward, implementing programs to achieve objectives which have little or no probablity of success due to a

for local consensus, understanding, and support by these most important groups of people.

The present study demonstrates the sensitivity and selectivity of one attitude measuring technique, Osgood's Semantic Differential, in providing critical information to decision-makers engaged in the establishment of institutional goals and objectives. If one assumes that, other things being equal, attitudes parallel and predict behavior, then attitudinal studies can provide useful information about the evolution of the decision making process and the probability of success of existing programs.

REFERENCES

Girod, G. R., ***Writing and assessing attitudinal objectives,*** Columbus: Charles Merrill, 1973, p. 55.

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H., ***The measurement of meaning,*** Urbana: University of Illinois Press, 1957.

Smith, R. G., Validation of the semantic differential, ***Speech Monographs,*** 1963, 30(3), 50-55.

Snider, J. R. and Osgood, C. E. (Eds.), ***The semantic differential technique,*** Chicago: Aldine, 1969.

APPENDIX I

Informal Criteria for Attitude Statements

1. Avoid statements that may be interpreted in more than one way.
2. Avoid statements that are irrelevant to the psychological object under consideration.
3. Avoid statements that are likely to be endorsed by almost everyone or by almost no one.
4. Select statements that are believed to cover the entire range of the affective scale of interest.
5. Keep the language of the statements simple, clear and direct.
6. Statements should be short, rarely exceeding 20 words.
7. Each statement should contain only one complete thought.
8. Statements (test questions - not the choices or answers) containing universals such as ***all, always, none,*** and ***never*** often introduce ambiguity and should be avoided.
9. Words such as ***only, just, merely,*** and others of a similar nature should be used with care and moderation in writing statements (test questions).
10. Whenever possible, statements should be used in the form of simple sentences rather than in the form of compound or complex sentences.
11. Avoid the use of words that may not be understood by those who are to be given the complete scale.
12. Avoid the use of double negatives.

Girod, G. R., ***Writing and Assessing Attitudinal Objectives,*** Columbus: Charles Merrill, 1973, p. 55.

APPENDIX II

The purpose of this study is to measure the meaning of some of the college-wide objectives which were distributed to all staff members. Please make your judgments on the basis of what these things mean to ***you.***

At the top of each page of this booklet you will find a different objective to be judged. Under each objective is a set of scales. Please rate each objective on each of these scales.

Here is how you are to use these scales.

If you feel that the objective at the top of the page is ***very closely related*** to one end of the scale, you should place your check-mark as follows:

fair X : ___ : ___ : ___ : ___ : ___ : ___ unfair

or

fair ___ : ___ : ___ : ___ : ___ : ___ : X unfair

If you feel that the objective is ***quite closely related*** to one or the other end of the scale (but not extremely), you should place your check-mark as follows:

strong ___ : _X_ : ___ : ___ : ___ : ___ : ___ weak

or

strong ___ : ___ : ___ : ___ : ___ : _X_ : ___ weak

If the objective seems ***only slightly related*** to one side as opposed to the other side (but is not really neutral), then you should check as follows:

active ___ : ___ : _X_ : ___ : ___ : ___ : ___ passive

or

active ___ : ___ : ___ : ___ : _X_ : ___ : ___ passive

If you consider the objective to be ***neutral*** on the scale, both sides of the scale ***equally associated*** with the objective, or if the scale is ***completely irrelevant***, unrelated to the objective, then you should place your check-mark in the middle space:

safe ___ : ___ : ___ : _X_ : ___ : ___ : ___ dangerous

IMPORTANT:

(1) Place your check-marks ***in the middle of spaces,*** not on the boundaries:

	This	Not This
	: ___ : ___ : _X_ :	___X___ :

(2) Be sure you check every scale for every objective - ***do no omit any.***

(3) Never put more than one check-mark on a single scale.

(4) Make each item a separate and independent judgment.

Work rapidly. Do not worry about individual items. It is your first impressions, your immediate feelings about the items, that are important.

The purpose of the college is to assist all members of the college staff to develop the best of their human potential by: helping them develop their capacity for critical thinking.

1.	Hard	___ : ___ : ___ : ___ : ___ : ___ : ___ Soft
2.	Foolish	___ : ___ : ___ : ___ : ___ : ___ : ___ Wise
3.	Honest	___ : ___ : ___ : ___ : ___ : ___ : ___ Dishonest
4.	Bad	___ : ___ : ___ : ___ : ___ : ___ : ___ Good
5.	Valuable	___ : ___ : ___ : ___ : ___ : ___ : ___ Worthless
6.	Unbelievable	___ : ___ : ___ : ___ : ___ : ___ : ___ Believable
7.	Strong	___ : ___ : ___ : ___ : ___ : ___ : ___ Weak
8.	Unsuccessful	___ : ___ : ___ : ___ : ___ : ___ : ___ Successful
9.	Fair	___ : ___ : ___ : ___ : ___ : ___ : ___ Unfair
10.	Passive	___ : ___ : ___ : ___ : ___ : ___ : ___ Active

PART 4
Evaluation of Support Department Objectives

Pam Swinford

I. Introduction

A department that wants to evaluate the effectiveness of its services by finding out if it is accomplishing its objectives may rely on informal feedback from other departments and individuals. A more accurate evaluation, however, involves the use of measurements and indicators.

In the context of the Participative Management Project, a **measurement** is defined as any behaviors or data which ***directly*** identify effectiveness in achieving an objective. An **indicator** is defined as any behaviors or data which ***indirectly*** imply progress in achievement. In other words, measurements assess the objective itself, while indicators assess conditions, events, and outcomes logically related to the objective.

The evaluation of an objective depends in large part on how the objective is written. Generally the less specific an objective, the more that indicators will have to be used as verification of accomplishment. To derive a measurement, activities must be quantifiable and numbers assigned so that comparisons can be made between the desired objective and the actual accomplishment.

Each department has the option of deciding which objectives can be specific, and therefore measurable , and which cannot, or need not, be specific. The following guidelines are offered to help departments make that decision by showing how an objective can be measured.

II. Writing a Measurable Objective

The first step is ***to state it clearly***. A measurable objective:

1. focuses on results;
2. is realistic and attainable;
3. establishes a quantitative standard to be achieved;
4. defines an "audience" or those at whom service is directed;
5. occurs within a specific period of time.

1. Focus on results

Initial attempts at writing objectives often produce statements that focus on activities rather than results. Activity-oriented statements describe functions rather than outcomes. Some examples of activity-oriented statements are:

To ***develop*** recruitment materials for older adults.
To ***promote*** the economical acquisition of equipment and supplies.
To ***serve*** as a resource agency.
To ***create or assist*** in the production of instructional and/or communications media.
To ***provide*** means of motivation and reinforcement of instructional classified staff.

The first example is an objective written by the community relations department at the administrative center. This statement and its subsequent revisions will serve as a model to demonstrate how an objective can be written so that elements leading to evaluation are "built in."

The simplest way to convert the focus of the statement ***To develop recruitment materials for older adults*** from an activity--the process of developing materials--to a result is to ask ***why*** the activity is being done. The answer will produce a statement that begins to focus on results:
1st revision

To recruit older adults as students through direct mail materials.

2. Make it realistic and attainable

Once the focus is on results, it is important that those results are realistic. An objective can miss the mark in one of two ways: It can be written so narrowly that little energy is needed to attain it; or the objective can be written so that it is impossible to reach.

The statement ***To recruit older adults as students through direct mail materials*** is too general for assessment. It implies that all recruitment of older adults can be accomplished through direct mail materials, which is unrealistic. A direct mail brochure is merely support for the work of college admissions personnel, counselors, etc. But, a brochure ***can*** carry appropriate information and encourage a positive response. The following statement is a realistic revision:
2nd revision

To gain a response to a direct mail recruitment brochure aimed at older adults.

3. Decide on a quantitative standard

This is the point at which activities are translated into numbers to aid in measurement. The numbers should represent a standard of achievement *desired* for the objective. The statement ***To gain a response to a direct mail recruitment brochure aimed at older adults*** is still vague because the degree of response is not indicated and there is no base for measurement.

Direct-mail marketing theory maintains that a 5% response is exceptionally good. With this background knowledge, the community relations department needs to determine a percentage of response that will provide quantitative evidence of the effectiveness of the brochures. The following revision might satisfy these requirements

3rd revision

To gain a 4% response to a direct mail recruitment brochure aimed at older adults.

4. Define the audience

Support departments often have problems defining their audience. That audience could be as large as all potential students in a community, or as small as office co-workers. Very large groups must be defined or broken into manageable parts so that activities can be focused and results can be measured. The older adults in the statement ***To gain a 4% response to a direct mail recruitment brochure aimed at older adults*** need further definition. The following revision specifies the group:

4th revision

To gain a 4% response to a direct mail recruitment brochure aimed at ***5,000 adults 60 years of age and older.***

5. Set a deadline

The only element missing from the last revision is a deadline. The work of support departments is not always easily organized according to semesters. Projects may begin and end at any time, and evaluation may be an ongoing activity. Unless a deadline is set for achievement of a goal, implementation could be delayed. The following revision incorporates a deadline:

final revision

To gain a 4% response ***by August 1*** to a direct mail recruitment brochure aimed at 5,000 adults 60 years of age and older.

III. Mechanics of Measurement

Once an objective is properly written, the mechanics of measurement need not be difficult. There are probably as many ways to derive indicators and measurements as there are support activities within the institution. The three types of evaluation, described on page 2 of this handbook are internal data, surveys and tests. Support departments probably will use internal data most frequently, surveys occasionally and tests rarely in evaluation. When data for measurements must be derived from the audience, "capturing" that audience could be a problem because of size, mobility and infrequency of direct contact. For this reason, evaluatory instruments may have to be built into administrative processes

To derive a measurement for the objective ***To gain a 4% response by August 1 to a direct mail recruitment brochure aimed at 5,000 adults 60 years of age or older***, the following procedures could be used:

1. Include a self-addressed, postage-paid response card in the brochure. Count the returns.
2. Instruct the reader to call a special extension, i.e. extension A. Count the number of callers who ask for that extension.
3. Compare the number of these responses to the total number of brochures mailed to derive the percentage of response.

The methods of measurement are limited only by a department's energy and available time. In most cases, if the statement of objective incorporates the five requirements for measurability--it focuses on results, is realistic, includes a quantitative standard, defines the audience and sets a deadline--the measurement is a logical outcome, rather than a mathematical mystery.