

Program Evaluation—A Broader Definition

Sara M. Steele

The meaning of evaluation is often overlooked in doing program evaluation. Our view of evaluation may be too small. The literature of program evaluation is changing rapidly with broader frameworks emerging. This article describes the three essential elements in evaluation—criteria, evidence, and judgment. The author says, "Evaluation must be purposeful and not done just for its own sake. Evaluation should contribute to the present program or to further and future programs."

Having trouble getting around to program evaluation? Does it frustrate you? One of the reasons we sometimes have difficulty with evaluation is that our concept of evaluation may be so small that we feel hemmed in. For years Extension has used definitions of Extension educational evaluation like:

The process of determining the change in behavior of people resulting from extension educational programs.¹

or

Evaluation is the process of determining the extent to which objectives have been attained.²

These definitions have guided major Extension program evalua-

tions such as: Extension practice adoption studies of the 1940s,³ the evaluation of the Farm and Home Management Program in the 1950s,⁴ the evaluation of work with low-income families in the 1960s,⁵ the consumer marketing program evaluations that have been carried on during the past 20 years,⁶ and many smaller evaluations done by individuals on their own programs.

Much of the framework for this concept of evaluation was drawn from the Tyler approach to curriculum development.⁷ Much Extension literature has focused primarily on methodology involved in collecting evidence of behavioral change.⁸ This evidence is important to administration in justifying pro-

SARA M. STEELE is Associate Professor, Department of Agricultural and Extension Education, and Coordinator, Staff Development, University Extension, University of Wisconsin, Madison, Wisconsin.

grams and securing additional budget for Extension. Unconsciously, many Extension professionals have accepted this concept of evaluation as being completely synonymous with a concept of program evaluation. To them, program evaluation is only determining the results of a program.

Recently there has been a major expansion both in the quality and quantity of the literature dealing with program evaluation in other educational fields.⁹ Federally funded educational and social action programs are required to include evaluation. At least three major universities have developed evaluation centers.¹⁰ The literature from other fields now suggests that the Extension concept of evaluation, *although still important*, may be only a fraction of the total concept of program evaluation needed.

This article explores changes in concept in terms of definition, elements that need further development, and characteristics of the program that need to be examined.

Contemporary Definitions of Evaluation

A concept is much more than a definition. However, definitions are handy starting places for analyzing the state of our understanding of a concept.

Two recent definitions help to set a broader base for a concept of program evaluation which may be important to Extension. The first focuses on the purpose of evaluation:

Generally, evaluation means the provision of information through formal means, such as criteria, measurement, and statistics, to serve as rational bases for making judgments in decision situations. To clarify this definition, it will be useful to define several key terms. A decision is a choice among alternatives. Judgment is the assignment of values to alternatives. A criterion is a rule by which values are assigned to alternatives, and optimally such a rule includes the specification of variables for measurement and standards for use in judging that which is measured. . . . Stated simply, evaluation is the science of providing information for decision-making.¹¹

The second re-emphasizes these elements a little differently:

Evaluation is the systematic process of judging the worth, desirability, effectiveness, or adequacy of something according to definite criteria and purposes. The judgment is based upon a careful comparison of observation data with criteria standards. Precise definitions of what is to be appraised, clearly-stated purposes, specific standards for the criteria traits, accurate observations and measurements, and logical conclusions are the hallmarks of valid evaluation.¹²

Essential Ideas in a Concept of Evaluation¹³

There are two important ideas to be drawn from these definitions.

First, evaluation must be purposeful and not done just for its own sake. Evaluation should contribute to the present program or to further and future programs.

Although there are many contributions that evaluation can make to education and the educator, its most important role is that of contributing to decisions about the program while it's in process or to decisions about further and future programming. Evaluation should be a part of answering questions about the appropriateness of methods, content, and program approach so that Extension programs can continually be improved. Evaluation should be an input to the programming process rather than an end to be achieved in that process.

Secondly, evaluation has three essential elements—criteria, evidence, and judgment.

Evaluation doesn't occur unless all three of these function. There must be criteria against which the program is judged, evidence of the extent to which the program meets those criteria, and a judgment of the extent to which the criteria were met.

To illustrate the roles of each of these three elements, suppose you're evaluating a new type of pencil to determine what kind your office should purchase. Element 1: You list the things you want in a pencil. In other words, you'd establish some criteria. You might say that you expect the pencil to:

1. Produce a mark of specific darkness.

2. Produce that mark with only a given degree of pressure.
3. Be smooth and fit your hand easily.
4. Cost less than three cents each when purchased in quantity.

Element 2: By writing with samples of the new pencils and examining the heaviness of the mark left on the paper, considering the amount of pressure you exert to leave the mark, considering the feel of it in your hand, and looking at the cost per pencil, you collect your evidence of the degree to which the pencil fits the criteria you've established.

Element 3: You then make a judgment about each of the criteria. Is the mark too light? Do you have to press too hard? Is it comfortable in your hand? Is it within the price range you've set? And finally you make your overall judgment—yes, this is an excellent pencil; we should place our order with this company.

On the other hand, let's say you've just collected evidence both from what the salesman said and from what you observed, but you didn't think much about your particular needs. You'd probably buy the pencil because the salesman went to school with your brother or some other irrelevant reason rather than because you judged that it met your needs.

Judging a pencil is much simpler than judging a program. Yet the point is that with only description we often are unable to really use the evidence in a way that's meaningful in program decisions.

Too often what we have called program evaluation is merely program description. Although there is evidence of results, there is no indication either of criteria or of judgment of the program which uses the evidence and results in meaningful conclusions.

Based on these three essential elements, we suggest this definition of evaluation:

Program evaluation is the process of judging (or a *judgment* as to) the worth or value of a program. This judgment is formed by comparing *evidence* as to what the program "is" with *criteria* as to what the program "should be."¹⁴

A third important idea related to the first two is: The programming decision that needs to be made identifies the focus of the evaluation. Programs can be judged (criteria stated and evidence examined) in terms of several different characteristics. How the evaluation will be used should determine the characteristic or characteristics that will be examined.

This idea will be discussed later. But first, let's look more closely at the three essential elements in evaluation and the extent to which they need further development as a part of a concept of program evaluation in Extension.

Evidence

Extension has made considerable progress in its understanding of evidence collection. As the number of

Extension staff members holding advanced degrees rises, the number having had at least one course in research methodology and statistics will also increase. In addition, the number of good references on how to develop questionnaires, tests, and other instruments for collecting evidence has multiplied in the last few years.¹⁵ The greatest need now, in terms of evidence collection, is for expert consultant help in developing sound instruments for measuring a specific kind of change such as change in attitude or knowledge gain. There is also a great need for computer and similar kinds of help that will handle the work of tabulating data and calculating statistics.

Much of our understanding of evidence-providing procedures focuses on summative evaluation (that done after a program is completed). We must give more attention to how we instantaneously collect and analyze evidence while the program is actually in progress and to how we use these data to improve and alter the original program design.

The areas where Extension in general and many individual Extension workers in particular need to expand their concept of evaluation are those of criteria and judgment. The ability to use these two elements successfully is important in all forms of evaluation and will greatly strengthen casual and semi-systematic evaluation.

Criteria

A criterion is something against

which something else can be judged. It may be a rule, a standard, a norm, or an object, condition, or behavior that is considered "good" or "ideal." It's a description or image of what a valuable (suitable, high quality, effective, important, and/or efficient) program is like.¹⁶

Programs are seldom judged on one criterion. Usually several individual criteria are developed into sets of criteria. Often both *macro*- and *micro*-level criteria are needed. Micro-level criteria set the standards for making specific judgments of parts of the program or describe subparts of a more complex standard. Performance criteria for determining when objectives have been met are micro-criteria. They help you decide the extent to which the individual objective has been accomplished.

For example, let's say a program has an objective for individuals to increase their participation in continued learning. How would you determine if this objective were attained? Your micro-criteria would state some specific things which the person would do which would show that he is indeed increasing his participation in learning. You might list criteria like:

1. Reads at least three more nonfiction books than he did last year.
2. Enrolls in at least one more class, workshop, or seminar that he did last year.
3. Watches educational programs on television more often than last year.

There are several other criteria (micro as used in this example) that would describe behavior and changes in behavior related to participating in continuing education. The evaluator must identify the range in appropriate criteria, and which ones are most relevant to his purpose. He'll have to make other decisions about how the criteria are to be used. Out of the various forms that such participation can take, does he expect the participant to show a change in behavior on all criteria? Will he accept an increase in any one of them? Must it be at least half? Is one more crucial than the other? If so, he might say that activity must have been increased related to Criterion No. 1 and to any other two criteria. Back to the pencil example. Cost may be the most important criterion. You might accept a little less quality in the other three criteria if the cost were right.

Macro-criteria help you add all of the smaller judgments into an overall judgment. For example, you might say, "We'll consider this a very successful program if 50 per cent of the participants adopt at least 5 out of the 8 practices." Or, "We'll consider this an outstanding program if there's evidence that at least three of the objectives have been met and an important change has occurred in the community as a result."

Where do we get criteria for judging programs? Some criteria are developed from principles or "basic truths." Some are developed from

theory or fairly well-accepted ideas. Others are developed from past experience and individual philosophy.

Criteria, in the sense of norms, can be built up as we build a pool of information about similar programs carried on under different conditions. This is one reason why sharing the results of evaluation through written reports is so essential.

In some instances, there are sets of criteria, developed by "experts," that can be adapted to our situation. For example, there are many sets of criteria for judging a project leader's performance. Here's an example of one that's set up in a form for leader self-evaluation:

- is enthusiastic about the job and the subject to be taught.
- puts ideas from the project into practice at home.
- has a definite purpose and directs teaching toward this purpose.
- keeps participation moving and on the track.¹⁷

In other instances, groups of colleagues have worked out a set of criteria for their particular programs. Often the programmer has to develop a set of criteria applicable to the decisions he must make.

Statements of criteria should be practical, relevant, clear, accurate, directive, and educational. They must be normative in the sense that they are stated so a judgment can be made. It's as important that the criteria be valid, reliable, and objective as it is that the evidence be so.

We usually focus on the degree or extent to which a program is suc-

cessful. Criteria, therefore, usually are developed so that differences in achievement can be identified along a scale or continuum. We then judge the extent to which we feel the program measures up.

Criteria come in all shapes and sizes. To be usable, they must be precise enough so sound judgments can be made. The sets of criteria used in evaluation must be stated very specifically and be concrete enough that others have the same understanding of the condition or quality that must be present for the program to be judged as valuable. For example, rather than saying, "He should prune a grapevine correctly," the criteria should specify where and when he must cut to be correct.

It's not easy to select and refine criteria. Many criteria can be applied to programs. The programmer must select or develop the set that will be of most use to him in getting the answers that will help him make decisions. When his idea of a standard is hazy or broad or nebulous, he must force himself to work with the idea until he can break it into constituent parts that can easily be identified and easily examined to see whether they are present in the program. Presently, many of the criteria we use in evaluating programs are within our subconscious. We use them without really crystalizing them to the extent where we can examine their soundness.

At a broader level, there must be more thought given to such questions as: When is an Extension pro-

gram considered successful? When does a program reach its maximum value? To what extent is success defined in terms of the situation?

For example, is there a difference in the interpretation of a 20 per cent change in practice adoption when it occurs at the beginning of the introduction of a new practice compared to when it occurs at the middle of the adoption cycle? Should there be a difference in expectations when the program deals with material that's difficult to grasp or is contrary to current thinking and habit? What is realistic to expect in terms of knowledge gained for a limited amount of input; for example, a two-hour meeting, a ten-minute radio broadcast?

The process of selecting and refining criteria, although frustrating, is one of the most important educational experiences in programing. When you know exactly what you aim to accomplish and what defines quality in programing, the teaching task is easier and resources are generally used to better advantage. Knowledge and good use of appropriate standards leads to better programs.

Criteria are established by many people. The programmer must take into consideration the criteria generally held by his clientele, the potential clientele who failed to participate, the general public and the legislators who represent them, experts in a given field, and the Extension agency. The criteria that the programmer actually applies may well be a blend of criteria from these

various sources.

It's sometimes well to involve some of the relevant publics in the actual formulation and application of program criteria. Too often we use program-planning groups, but fail to bring them back as program evaluators. Bringing the planners together to react to the program as designed and then as completed can be an important part of the program evaluation. It's often valuable to have representative program participants examine and interpret evidence that has been collected for a program. They can assist the programmer in making the final judgments as well as adding their own judgments to the evidence that's finally used.

It's clear from the preceding discussion of criteria that one's judgment as to what criteria will be used and how they'll be used greatly affects the final judgments that will be made about the program.

Judgment

We can't dodge responsibility for judgment. We can evade it by simply gathering evidence with little rationale other than that it describes the program. But if we do that we're describing and not evaluating.

Judgment is essential in program evaluation. Judgment has been described as a cognitive process with the following characteristics:

1. The main inputs to the process, that which is to be judged, are given and available; obtaining, discovering,

disagree

- or formulating them is not part of judgment.
2. The domain of the output—the set of admissible responses—is simple and well-defined prior to the judgment.
 3. The process is not a simple transduction of information; judgment adds information to the output.
 4. The process is not simply a calculation, or the application of a given rule.
 5. The process concludes, or occurs at the conclusion of, a more extended process.
 6. The process is rather immediate, not being extended in time with phases, stages, sub-processes, etc. (If such occur, they tend to be referred to as preparation for judgment.)
 7. The process is to be distinguished from searching, discovering or creating, on the one hand; and from musing, browsing, or idly obtaining on the other hand.¹⁸

And yet for most of us judgment is an intangible process—one that is hard for us to define and examine. We must know a good deal more about judgment as it occurs in program evaluation. Evidence may be valid, reliable, objective, and all the rest, but the interpretation of that evidence may be biased and the resulting judgment faulty. It's as important that the programmer strive for valid, reliable, and objective judgment as for scientific technique in evidence collection. We must accept the fact that judgments are subjective and work toward making them as objective as possible.

Judgments are made by people and are dependent on them. Judgments are influenced by the past experience and values of the individuals making them. Judgment is improved by experience—making judgments and testing them for their soundness. Professional judgment is one of the greatest assets of the professional. The testing of judgment through evaluation can be an important step in professional development.

The authors of the *Taxonomy of Educational Objectives II* say:

Although it is recognized that an individual is, on many grounds, entitled to his own opinion as well as his own judgments about the value of specific ideas, objects, or activities, one major purpose of education is to broaden the foundation on which judgments are based. Thus, it is anticipated that as a result of educational procedures individuals will take into consideration the greater variety of facets of the phenomena to be evaluated and that they will have in mind a clearer view of the criteria and frames of reference being used in the evaluation.¹⁹

And of the three elements in evaluation, judgment is clearly the most important. Judgment is the heart of the casual, everyday kinds of evaluations on which many of our most important program decisions are based. The most rigorous program evaluations in terms of amount of evidence haven't always been used in program decisions while many major decisions have

been made using very casual evaluation. The greatest challenge to the Extension evaluator is that of improving his judgment. Certainly, knowing how to use criteria and evidence will improve that judgment, but there's more to judgment than just having criteria and evidence. The evaluator must be able to use both in coming to sound conclusions.

Judgments About What

The third key idea listed earlier dealt with the fact that a program can be judged in regard to many different characteristics. Just as a cow can be judged in terms of appearance, milk production, or breeding record, so a program can be judged in terms of different aspects. There are at least five program characteristics that are sometimes evaluated. These characteristics have been called different things and aren't completely separate. Each contributes to the other.

- *Quality*: How good was it? What was the quality of the content, learning activities, media, teacher's performance? How did people react to it?
- *Suitability*: Did it meet the needs and expectations of the participants? Was it at the appropriate level of difficulty? Did it meet the expectations of the community? Was it within the mission of the programing unit?
- *Effectiveness*: What did it accomplish? How well did it accomplish its objectives?

- *Efficiency*: Were the accomplishments sufficient for the amount of resources required from the agency and the participants? Was this the best use of resources?
- *Importance*: How valuable was it to those who participated and to society? Was its importance sufficient to the resources that were involved?

Our old concept of evaluation focuses primarily on program effectiveness or how well the objectives were accomplished. Actually, however, much of our evaluation has been in terms of quality and suitability of program as judged by participant response to end-of-meeting sheets which explore such things as how the participant would rate the program, whether he thought he understood the content, whether the program was timed well, etc. Those holding to a definition of evaluation that includes only examining behavioral results, have in effect, said over the years that these end-of-meeting reactions weren't evaluation. However, if one accepts the fact that a program may be judged in terms of more than one characteristic, then these sheets contribute to evaluation, but explore different characteristics (quality and suitability) of the program.

There is value in making judgments about the quality and suitability of the program, and its effectiveness. It is also essential that contemporary educational programs be judged on still other characteristics.

When resources are limited, perhaps the most important charac-

teristics for examination are those of *efficiency* and *importance*. Both depend on the program's effectiveness, but consider more than effectiveness. We need to give more consideration to whether we're using our resources to best advantage and to the relative importance of programs, selecting those for major emphasis that will make the greatest contribution to the clientele and to society as a whole. Extension needs to give a good deal more attention to judging the relative importance of its many program opportunities.

Judgments as to potential importance must be made about the program design before it's actually implemented. We need more evaluation of programs when they're in the design stage, but our traditional concept of evaluation focuses on the extent to which an objective has been reached. It doesn't challenge the objectives. Sometimes those objectives should have been evaluated and altered before the program was launched. Achievement of a poor objective doesn't result in a good program.

Regardless of the point in programming at which a program is evaluated or the characteristics of the program being examined, the basic structure of evaluation—criteria, evidence, and judgment—applies. We need to: (1) develop criteria related to the different characteristics and understand what characteristic of a program we're actually using as the basis of judgment when we use a particular kind of criteria, (2) identify which criteria

are most important in which situations, and (3) increase our experience in making judgments of various characteristics and the quality of those judgments. Evaluation isn't necessarily easy, but a broad understanding can make it easier.

If we're to meet the challenge given in *A People and a Spirit*,²⁰ that of doing interpretive or explanatory evaluation (not just judging programs, but explaining why programs have or haven't been valuable), we probably will need to explore not only how programs measure up on all five of these characteristics but how the degree of "goodness" on each affects the other and the total value of the program.

Summary

We may be more effective in evaluation if we broaden our concept of evaluation by accepting the legitimacy of evaluating programs in terms of other characteristics in addition to effectiveness (that is, accomplishment of objectives). A concept of evaluation must include the subconcepts of criteria and of judgment as well as the subconcept of evidence.

Extension's greatest challenge in improving evaluation rests with our ability to improve criteria and improve our use of criteria in arriving at the kinds of judgments that will help us improve future Extension programs. Helpful literature in the field of program evaluation is increasing. There's much to be learned both from reading and from

actual everyday evaluative experience. We need more exchange of theory and of findings as we evaluate Extension programs.

Footnotes

1. Laurel K. Sabrosky, "Evaluation," in H. C. Sanders, ed., *The Cooperative Extension Service* (Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1966), p. 339.
2. Wilson Thiede, "Evaluation and Adult Education," in Gale Jensen, A. A. Liveright, Wilbur Hallenbeck, eds., *Adult Education* (Washington, D.C.: Adult Education Association of the U.S.A., 1964), p. 291.
3. Several such studies are summarized in H. C. Wilson and Gladys Gallup, *Extension Teaching Methods*, Extension Service Circular No. 495 (Washington, D.C.: USDA, Federal Extension Service, 1955).
4. Systematic evaluation was conducted in several states including New York, North Carolina, Missouri, and Wisconsin. See Donald E. Johnson and E. A. Wilkening, *Five Years of Farm and Home Development in Wisconsin: A Comparison of Participation and Control Families*, Research Bulletin No. 228 (Madison, Wisconsin: Wisconsin Cooperative Extension Service, 1961); C. D. Marsh, *An Evaluation of the Farm and Home Development Approach to Agricultural Extension Work in North Carolina*, North Carolina Extension Evaluation Studies No. III (Raleigh, North Carolina: North Carolina State College, 1962); and Frank Alexander and James W. Longest, series of reports evaluating the farm management program in New York State, Ithaca, New York, Extension Studies Office, n.d.
5. Such studies were carried out in Boston, Kansas City, and San Antonio. See Starley Hunter *et al.*, *Families in an Urban Enclave* (Columbia, Missouri: University of Missouri, Extension Division, 1965) and Starley Hunter *et al.*, *The Families and Their Learning Situations, South End Housing Development, Boston, Massachusetts* (Amherst, Massachusetts: University of Massachusetts, Cooperative Extension Service, 1965).
6. Among the many are Mescal Johnston and Ward F. Porter, *Evaluation of the Little Rock Marketing Information Project*, Bulletin No. MP88 (Fayetteville, Arkansas: University of Arkansas, Agricultural Extension Service, n.d.) and Ward F. Porter *et al.*, *Food Buying, Knowledge-Concerns, Practices*, Agricultural Experiment Station Bulletin No. 456 (Morgantown, West Virginia: West Virginia University, 1961).
7. Ralph W. Tyler, *Basic Principles of Curriculum and Instruction* (Chicago, Illinois: The University of Chicago Press, 1950), pp. 68-81.
8. Darcie Byrn, ed., *Evaluation in Extension* (Topeka, Kansas: H. M. Ives and Sons, 1959).
9. Recent literature includes such books and articles as Ralph W. Tyler, ed., *Educational Evaluation: New Roles and New Means* (Chicago, Illinois: The University of Chicago Press, National Society

- for the Study of Education, 1969); Edward A. Suchman, *Evaluative Research* (New York, New York: Russell Sage Foundation, 1967); Fred T. Wilhelm, ed., *Evaluation as Feedback and Guide* (Washington, D.C.: NEA, Association for Supervision and Curriculum Development, 1967); R. E. Stake, ed., *Perspectives of Curriculum Evaluation* (Chicago, Illinois: Rand McNally and Company, 1967); R. E. Stake, "The Countenance of Educational Evaluation," *Teachers College Record*, LXVIII (April, 1967), 523-40; Egon Gruba, "The Failure of Educational Evaluation," *Educational Technology*, IX (May, 1969), 29-38; and Thomas R. Owens, "Suggested Tasks and Roles of Evaluation Specialists in Education," *Educational Technology*, VIII (November 30, 1968), 4-9.
10. Evaluation centers have evolved at such major universities as Evaluation Center, The Ohio State University, College of Education, Columbus, Ohio; Center for Instructional Research and Curriculum Evaluation, University of Illinois, Urbana, Illinois; and Center for the Study of Evaluation, University of California, Los Angeles, California.
 11. Daniel L. Stufflebeam, "Toward a Science of Educational Evaluation," *Educational Technology*, VIII (July 30, 1968), 5-12.
 12. Wilbur Harris, "The Nature and Function of Educational Evaluations," *Peabody Journal of Education*, XLVI (September, 1968), 95-99.
 13. Sara M. Steele, *Developing a Concept of Program Evaluation* (Madison, Wisconsin: National University Extension Center, 1970).
 14. *Ibid.*
 15. Examples of books on methodology relevant to evaluation include Edward J. Furst, *Constructing Evaluation Instruments* (New York, New York: Longmans, Green and Company, 1958); Robert Ladd Thorndike and Elizabeth P. Hagen, *Measurement and Evaluation in Psychology and Education* (New York, New York: John Wiley & Sons, Inc., 1962); Claire Selltitz *et al.*, *Research Methods in Social Relations* (New York, New York: Holt, Rinehart and Winston, Inc., 1964); Leon Festinger and Daniel Katz, *Research Methods in the Behavioral Sciences* (New York, New York: Holt, Rinehart and Winston, Inc., 1965); Deabold B. Van Dalen, *Understanding Educational Research* (New York, New York: McGraw-Hill Book Company, 1966); Frederick Williams, *Reasoning with Statistics* (New York, New York: Holt, Rinehart and Winston, Inc., 1968); and Charles H. Backstrom and Gerald D. Hursch, *Survey Research* (Evanston, Illinois: Northwestern University Press, 1963).
 16. For references related to criteria, see David Ryans, "Notes on the Criterion Problem in Research, with Special Reference to the Study of Teacher Characteristics," *The Journal of Genetic Psychology*, XCI (1957), 33-61; Alexander W. Astin, "Criterion Centered Research," *Educational and Psychological Measurement*, XXIV (No. 4, 1964), 807-22;

- H. E. Brogden and E. K. Taylor, "The Theory and Classification of Criterion Bias," *Educational and Psychological Measurement*, X (1950), 159-86; and Sara M. Steele and Gwenna M. Moss, "The Criteria Problem in Program Evaluation," paper presented at the Adult Education Research Conference, Minneapolis, Minnesota, 1970.
17. "Qualifications of a Good Project Leader," *Leads for Leaders*, Vol. I, No. 8a (Madison, Wisconsin: University of Wisconsin, n.d.), mimeographed.
 18. Benjamin Kleinmuntz, ed., *Formal Representation of Human Judgment* (New York, New York: John Wiley & Sons, Inc., 1968), pp. 5-6.
 19. Benjamin S. Bloom, ed., *Taxonomy of Educational Objectives: Handbook I* (New York, New York: David McKay Company, Inc., 1956), p. 186.
 20. USDA/NASULGC Extension Study Committee, *A People and a Spirit* (Fort Collins, Colorado: Colorado State University, 1968), p. 42.