DOCUMENT RESUME

ED 037 188 | 24 | HE 001 450

AUTHOR	Pace, C. Robert
TITLE	An Evaluation of Higher Education: Plans and Perspectives.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
REPORT NO	CSE-51
BUREAU NO	BR-6-1646
PUB DATE	Jan 69
NOTE	15p.
EDRS PRICE	EDRS Price MF-$0.25 HC-$0.85
DESCRIPTORS	*Educational Philosophy, *Evaluation, *Evaluation Criteria, Evaluation Methods, Evaluation Needs, *Higher Education, *Investigations
IDENTIFIERS	*Center for the Study of Evaluation

ABSTRACT

This report presents the philosophy of the Center for the Study of Evaluation at UCLA. It points out that any evaluation of higher education in America must deal with 3 fundamental problems: 1) diversity of objectives and functions across the full range of institutions of higher education; 2) differences between college environments; and 3) range of student abilities. A more subtle obstacle to effective evaluation has been the traditional concept of what evaluation actually is. Traditionally, institutional goals have been the object of evaluation, but it is more important to look at the consequences of particular politics or programs. Although "control" and "focus" have characterized the style of inquiry in the past, the investigator needs to be adventurous and explorative in order to evaluate changing social and educational developments. "Exploration" involves searching, probing, and testing alternatives and interactions, but it also connotes a freedom to seek new methods and values ingenuity. A new model for evaluation programs sees the role of the evaluator as a social scientist rather than teacher, missionary, or reformer. Given this concept of evaluation, it is necessary to think about the content of evaluation in a broadly inclusive rather than restrictive fashion. The remainder of the report consists of a description of the methodological approaches of the Center. (DS)

BR-6-1646
CSE R#51
PA-24
OE-BR

ED037188

AN EVALUATION OF HIGHER EDUCATION: PLANS AND PERSPECTIVES

C. Robert Pace

CSE Report No. 51

January 1969

Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California

HE 001 450

CENTER FOR THE STUDY OF EVALUATION

Marvin C. Alkin, Director
Publications Committee:
James W. Trent, Chairman
Theodore R. Husek
Sherman J. Pearl
Audrey Schwartz

UCLA Graduate School of Education

The CENTER FOR THE STUDY OF EVALUATION is one of nine centers for educational research and development sponsored by the United States Department of Health, Education and Welfare, Office of Education. The research and development reported herein was performed pursuant to a contract with the U.S.O.E. under the provisions of the Cooperative Research Program.

Established at UCLA in June, 1966, CSE is devoted exclusively to finding new theories and methods of analyzing educational systems and programs and gauging their effects.

The Center serves its unique functions with an inter-disciplinary staff whose specialties combine for a broad, versatile approach to the complex problems of evaluation. Study projects are conducted in three major program areas: Evaluation of Instructional Programs, Evaluation of Educational Systems, and Evaluation Theory and Methodology.

This publication is one of many produced by the Center toward its goals. Information on CSE and its publications may be obtained by writing:

Office of Dissemination
Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California 90024

PREFACE

The main purpose of the present report is to describe the scope of an evaluation of higher education that is currently under way in the Center for the Study of Evaluation. By way of introducing the content of this study, we also have presented some thoughts about the concept and practice of evaluation in education.

We make no claim to having solved any problems. Instead, we have sought, by the very scope of our study, to create problems of methodology and analysis which we and others must try to solve by dealing with the complexities that a comprehensive appraisal requires. We do not discuss these technical problems in the present report.

It should be clear to the reader that there are many aspects of higher education not considered in the studies we describe--for example, economic factors, administrative and organizational factors, classroom teaching, political influences, etc. We are very interested in these aspects, and we hope to consider them more explicitly in future studies. For the present, we have tried mainly to extend the range of criteria used in judging the impact of education on the behavior of students and former students and to extend the range of environmental conditions that may have a bearing on this impact. Many commentators have said that the most obvious characteristic of higher education in this country is diversity. If one grants some truth to this observation, then in his efforts to evaluate higher education, he must take this diversity into account--in the range of objectives one considers and in the variety of institutional settings one studies.

In the total program of the Center for the Study of Evaluation, a considerable range of evaluation problems is being studied--classroom learning, elementary school system, program budgeting and cost-benefit analysis, math curricula, and other topics. The variety of projects testifies to our belief that evaluation is a pervasive concern. The higher education project illustrates one type of problem and one set of views about it.

C. R. P.

AN EVALUATION OF HIGHER EDUCATION: PLANS AND PERSPECTIVES

Most evaluative studies of higher education have dealt with one or a limited number of aspects of higher education rather than with the total enterprise. To consider the total enterprise as the object of study may well be beyond the capacity and ingenuity of any group of investigators. Yet the more nearly one can approach such comprehensiveness, the more likely it is that his results will contribute to better judgments and perspectives.

With some utility, economists characterize the economic state of the nation and predict the probable consequences of such policies as lower taxes or higher interest rates. Anthropologists characterize whole cultures. Ecologists are increasingly concerned with trying to understand man's total habitat and the intricate balance between organisms and environments. Architects and urban planners are increasingly concerned with trying to develop total environments that will enhance the quality of life. In each case, the efforts have led to a new appreciation of the complexities and the interacting forces that must be explored--whether of the economy, the culture, the biological habitat, or the urban environment.

In evaluations of higher education programs, such complexity has been found in occasional studies of single institutions, but studies that encompass large number of institutions have typically dealt with a smaller number of variables. The purpose of some of the larger nationwide studies has been to describe a complex phenomenon and explain how it works--as in the selective distribution of students to colleges or occupations. In other studies the purpose has been to evaluate how well one type of objective is attained--such as changes in scores on personality tests, in scores on achievement tests, or in students' plans to attend graduate school--with efforts to explain what accounts for the performance. If one agrees that the training of scientific and professional personnel, the encouragement of high school students to attend college and college students to attend graduate school, the acquisition of knowledge in broad fields, and the development of personality are in certain ways all among the important goals and functions of higher education, then these inquiries, taken together, have encompassed a significant range of issues and outcomes. The serious questions to which educational evaluators have addressed themselves provide clear evidence that there is a growing awareness of complexity and a growing desire for more adequate evaluations. Examples of significant local and national studies are Reeves, 1933; Eckert, 1943; Davis, 1965; Thistlethwaite, 1968; Astin, 1968; Trent & Medsker, 1968; and Wolfle, 1954.

A comprehensive national evaluation of higher education must deal with at least three fundamental problems: first, the diversity of objectives and functions across the full range of higher institutions, necessitating a set of criterion or outcome measures that reasonably reflect and acknowledge this diversity; second, the differences between college environments, necessitating the use of suitable measures for describing and classifying some of the major ways in which educational treatments and experiences differ from one another; and third, the range of student abilities, backgrounds, interests, and aspirations, necessitating the use of test data and other information to identify what kinds of students experience what kinds of treatments and attain what kinds of results.

There are many published tests that could be used in such a comprehensive appraisal; but using enough of them to cover a broad range of criteria, contexts, and characteristics has not been feasible because such a practice would be too time-consuming and would require unusual administrative and financial resources. The obstacle of feasibility can be overcome by developing shorter tests and by taking advantage of information that already exists in published sources and from previous or concurrent studies. The evaluation program described later in this report represents one serious effort to develop feasible measures that will permit a more comprehensive evaluation.

A subtler obstacle has been the traditional concept of what evaluation is. In the field of education, evaluation has been typically regarded as being concerned with how well the specifically defined objectives of a program are attained or, in experimental designs, with determining whether or not one educational treatment produces results different from those of an alternative treatment or a control group. The requirements for such a model are rigorous: (a) clear objectives--behaviorally defined, (b) random assignment of subjects to different treatments, (c) clearly differentiated treatments, and (d) criterion measures explicitly related to the objectives. There is some recognized slippage in the model, however, for investigators are usually admonished to be on the lookout for "side effects" and for alternative explanations of their results. Nevertheless, the model has served education well and has led to many solidly documented conclusions, as a reading of the _Handbook of Research on Teaching_ (Gage, 1963) will confirm. The shortcomings of the model can be illustrated by some rhetorical questions: What does one do when not all the relevant objectives are manifested in directly observable specific individual behavior? What does one do about deliberately trying to measure effects that are not objectives of the program? What does one do when random assignment of subjects to treatments cannot be accomplished? What does one do when he lacks clearly differentiated treatments? Or when a treatment changes during the course of the study? The traditional evaluation model provides no answers to such questions. Perhaps the answer it does provide is illustrated by a _New Yorker_ cartoon that appeared when the automatic record changer first came on the market--showing records flying off in all directions with the caption "If it isn't Bach, the hell with it!"

One of the most hopeful developments in higher education today is the willingness of many colleges to try new programs or, in the image of the cartoon, to play new music. The book _Higher Education: Some Newer Developments_ (Baskin, 1965) supports this fact. There are periods of social change that correctly can be described as turbulent. We are living in one of them. One cannot make change stand still so that an evaluator can compare "experimental _vs_ control" treatments. If there is a course in Negro History, one cannot exclude students who want to take it. If new math is a good thing, one cannot prevent students from encountering it. Indeed, the whole notion of random assignment of subjects to treatments is usually not tolerated in the reality of society where the pressure is for equal opportunity and the best possible treatment for everyone. Such controlled conditions can be set up only in a laboratory that is removed from the moral and dynamic imperatives of the larger society, and the results of such laboratory experiments may or may not be reproducible in more realistic settings.

If educational evaluators are to deal effectively with large problems that cannot be simulated in a laboratory--problems such as the effectiveness of a total institution, of a class of institutions, or of higher education in the United States--then they need a concept of evaluation that is both more comprehensive and more flexible than the familiar experimental model. For this level of complexity and reality, the evaluator must ask different questions, proceed in a different style, and have a new view of his role and purpose.

The central question is not "What are the objectives?". The central question is "What are the consequences?" If one should ask himself the question, "What are the objectives of the United States in Vietnam?" , no doubt some answers would come to mind. But if instead one asked himself the question, "What are the consequences of the war in Vietnam?", a much greater range of inquiry would immediately be suggested and required. An analogous situation exists with any large scale social phenomenon, and higher education in the United States is one such phenomenon. What are the consequences of the existence of higher education--of the fact that there are some 2,400 colleges, universities, and junior colleges, that there are some 6,000,000 or more enrolled degree candidate students, and that more than half of the nation's high school graduates enter an institution of higher education? What if there were no higher education institutions? What then would be the consequences? It seems obvious that the range of one's inquiry is guided by the questions one asks. "Objectives" are a subheading under "consequences." Of course, one hopes that among the consequences are some which are intended as objectives, but looking at the extent to which objectives are achieved will not answer the larger question about consequences. Therefore, the first requirement for a new model of evaluation is to begin with the question "What are the consequences?".

The second necessary element in a new model is one which relates to the appropriate style of inquiry. An apt term for this is "exploration." Traditionally, the style of inquiry has been characterized by the words "control" and "focus." Exploration is a freer style--one which encourages hunches, is uncommitted, and seeks discovery. If the program one hopes to evaluate is continually changing in methods, materials, personnel, and subjects, it is not necessarily true that it cannot be evaluated. To the contrary, one may discover that programs which are being modified continually are more effective than programs that remain relatively static. The spirit of the evaluator should be adventurous. If only that which could be controlled or focused were evaluated, then a great many important educational and social developments would never be evaluated--at least not by "evaluators;" that would be a pity. To suggest that the style of the controlled experiment needs to be replaced by an exploratory style does not mean that one's approach should be any less scientific. Exploration involves searching, probing and testing alternatives and interactions. It can be tough-minded, rigorous, and theory-based. But the word exploration also connotes a freedom to look around, to seek new measures and methods, and to value ingenuity and curiosity.

The third element for a revised concept of evaluation is a new view of the role of the evaluator and the purpose of evaluation. Historically, many people have seen evaluation as an instrument of reform. The reason for evaluating any present activity or program was to improve it. As a result of this approach, the parties to the activity or program had to be involved in its evaluation because their very involvement would increase the likelihood that they would be willing to change in the light of the findings. Thus, the process of carrying out an evaluation--group participation and cooperation--was directly related to achieving the purpose of evaluation, namely, change and improvement. Implicitly, the role of the evaluator was in some respects the role of missionary and reformer. Another common view of the purpose of evaluation is that it exists to give feedback to the decision-makers (presumably teachers and administrators). In this case the role of the evaluator is a kind of staff officer to the practitioner. Neither of these views seems quite appropriate for evaluations that deal with large, complex social or educational programs such as higher education in the United States. A clearly identifiable number of individuals who can be called the decision-makers does not exist. In realpolitik reforms are brought about by persuasion, pressure, consensus, and conviction--not by evaluators. What evaluators produce may, of course, contribute to the decisions that are made. One hopes that they will; but whether or not evaluators contribute to such decisions is not the issue. Whether or not the evaluator is also the politican and thus carries out his evaluation in ways designed to generate certain convictions is the issue. A more appropriate concept of evaluation would define the role of the evaluator as that of a social scientist and the purpose of evaluation as that of providing more complex bases for informed judgment. When asked about the relation between social scientists and administrators, Harold Lasswell is reported to have said that the role of the social scientist is to complicate the task of the decision-makers. The evaluator, as a social scientist, should see this pursuit as his role, too.

A concept of evaluation, appropriate for the study of large and complex institutions, can be summarized briefly as follows:

1. It begins with the central question, "What are the consequences?" rather than with the more limiting question, "What are the objectives?"

2. Its style of inquiry is more aptly characterized by the word "exploration" than by the words "control" and "focus."

3. It sees the role of the evaluator as a social scientist rather than as a teacher, missionary, reformer, or staff officer to the practitioners.

4. Its purpose is to provide more complex bases for informed judgment.

Given this concept of evaluation, one should be persuaded to think about the content of evaluation in a broadly inclusive rather than restrictive fashion.

In June, 1966, the United States Office of Education established at UCLA a Research and Development Center for the Study of Evaluation. The Center has been concerned with the theory and practice of evaluation at all levels of education--elementary, secondary, and higher--and with programs at different levels of complexity--from classroom learning to large school systems. One of the activities of the Center is focused on evaluation in higher education. A description of this activity constitutes the remainder of this report. The long-term program is more ambitious than what we shall describe here, for its ultimate intent is to explore a broad range of social indicators and consequences of higher education--both positive and negative--of the sort suggested earlier in this report. But our exploration of such widely-ranged consequences has not yet progressed to the point where we feel that some publication is merited. In this report we shall describe one comprehensive study that is now under way--a study concerned primarly with a range of outcomes of higher education that may be seen in the behavior of students and adults exposed to the college experience, and with the institutional characteristics and individual experiences that may explain the extent to which different outcomes are achieved.

Growth in knowledge and understanding, intellectual skills and interests, heightened awareness, appreciation, values, attitudes, citizenship, moral sensitivity, and activities as producers and consumers of "the good life"--all these and more are commonly discussed among the objectives of higher education. Yet such a broad range of goals is seldom dealt with in systematic evaluations of higher education. More frequently, judgments about the quality of institutions are based on only one type of criterion--their assumed or measured intellectuality. This method of judgment may define the place of the institution in the system but only on one dimension, and it reveals nothing about the effectiveness of the institution. Many colleges and universities, which are necessarily low on an intellectuality dimension because of the nature of their student input, may be highly effective in other respects such as the amount of change produced in their students or the role of their graduates in community service and active citizenship. As we approach more or less universal higher education, evaluations based on a single criterion are not only increasingly inadequate, but they may lead to a distorted national perspective of the different roles and differential effectiveness of higher institutions.

Moreover, most evaluations have focused on individual measurement rather than on institutional measurement. Yet higher education in any organized sense takes place in some sort of institutional setting. Recent research (Pace, 1967) has clearly demonstrated that large differences exist among different institutional settings. With more than half of all the high school graduates in the country now going to college and distributing themselves among institutions which in many cases differ radically from one another, how much do we really know about the kind of impact these varied institutions have upon their students?

In any national appraisal of higher education, there are really three potential sources of distortion: first, an inadequate range of criterion variables;

second an inadequate range of contextual or environmental variables; and third, an inadequately representative population of individuals and institutions. In the studies we are conducting, we are attempting to minimize each of these potential sources of distortion.

First of all, our population includes approximately 100 colleges and universities from all parts of the country, large and small, public and private. Within this national assortment of educational institutions we have included selective examples of each of eight major categories, or types, as follows:

1. Highly selective liberal arts colleges, private, nonsectarian
2. Strongly denominational liberal arts colleges
3. General liberal arts colleges
4. Highly selective universities, public and private
5. General comprehensive universities, public and private
6. State colleges and other universities having less extensive graduate programs than general comprehensive universities.
7. Colleges having a major emphasis on teacher education
8. Colleges having a major emphasis on engineering and sciences

Within each of these colleges and universities, we are obtaining data from random samples of incoming freshman, upperclassmen, and alumni, with larger samples from the larger institutions and smaller samples from the smaller institutions. Alumni and upperclassmen are being tested in the winter and spring of 1969 and freshmen in the fall of 1969.

The basic content of the questionnaire developed for this evaluation study is the same for each of the three groups, except for minor changes in wording and a few additions or omissions of items that would be applicable to only one of the three groups. Our description of the content will refer to the alumni questionnaire.

The first section of the questionnaire consists of activity scales dealing with a broad range of involvement in contemporary society and culture. Each of the activity scales contains from nine to twelve items. The topics of these scales are as follows: community affairs, national and state politics, art, music, literature, drama, education, science, religion, inter-cultural affairs, and international affairs. The internal structure of each of the scales is the same in that each scale includes activities ranging from some which are relatively simple, commonplace, and easy to do to ones which involve increasing amounts of interest, time, and commitment. The instructions for each scale are to "check each statement that describes an activity you have engaged in during the past year." The following statements describe the scale for art activities:

During the past year:

___ I talked about art with my friends.
___ I read critiques or reviews of art shows or exhibits in the newspapers or magazines.
___ I visited an art gallery or art museum.

___ I attended an exhibition of contemporary painting or sculpture.
___ I read one or more books about art, artists, or art history.
___ I bought a painting or piece of sculpture.
___ I attended an art study group or workshop.
___ I contributed money or time in support of some activity related to art.
___ I did some creative painting or other art work myself.

The number of activities checked in each scale provides a measure of the amount and depth of one's participation and interest. The number of different scales in which one checks more than some minimal number of statements provides a measure of the breadth of one's interests and involvement. Moreover, since many of the scales have certain roughly parallel items, one can derive additional indexes that have some relevance to the consequences or effects of higher education. For example, one can obtain a score indicating the number of different fields or topics about which one talks or discusses with his friends, an index of the number of different fields in which one has read a book, and the number of different fields in which one attempts to keep up with current events through the newspapers, magazines, or television. With the arts (art, music, literature, and drama) three additional indexes can be derived: an index of the extent to which one exposes himself to a variety of opinions through the reading of reviews and criticisms, an index of one's exposure to contemporary works, and an index of engagement in personally self-expressive behavior. From the civic and political scales, one can obtain an index of activism. Thus, the activity scales provide a rich source of information relevant to many of the intended or potential outcomes of higher education. Moreover, some of the items in these scales are similar or identical to ones which were initially developed by the writer for alumni surveys at American University and Syracuse University in 1946-47 and which were also used in a nationwide survey of college graduates made in 1947-48 by the research division of Time magazine. The possibility of observing 20-year trends is an added benefit of significance.

The second section of the questionnaire is designed to provide two kinds of measures: first, a measure of knowledge and awareness about certain major changes that are taking place in American society and, second, a measure of attitudes toward such changes. Knowledge of what is and is not happening in society is fundamental to any intelligent adaptation to changing conditions and in this sense may be regarded as a culminating objective of liberal education. The primary sources for constructing items about major social changes have been various publications from the American Academy of Arts and Sciences, the National Research Council, and other governmental or independent bodies of social analysts. In range of content, the various items deal with changes in the labor market, the economy, the environment, education, science, government, industry, and other important aspects of the American scene. Here are a few sample statements in the test:

___ More people are coming to realize and accept the value of self-expression--for example, in the arts.
___ Scientists and professionals are having an increasingly important

influence on economic and governmental policies.

____ Increasingly, government is controlling the markets for the most advanced industries.

To each of these statements and some 25 others like them, two types of responses are solicited: first, a knowledge response indicating whether one thinks that the statement is or is not generally true, and second, an attitude response indicating whether one thinks the change described by the statement would be desirable or undesirable if it has in fact occurred or is occurring. The recognition of change and the readiness to deal with it can thus be compared with the denial of change and resistance to it.

The third section lists various educational objectives or potential benefits and asks the respondents to rate the extent to which the college experience was influential in relation to those objectives or benefits. In another section, we ask several questions about the existential value of college experience, apart from any instrumental or vocational benefits.

The sections of the questionnaire which we have described are concerned with criterion measures, that is, with activities and interests, knowledge, attitudes, judgments, and feelings related to a broad range of intended or relevant outcomes of higher education. The last two sections of the questionnaire deal with the nature of the school and college experience and with a variety of personal information that is intended to throw light on some of the individual and environmental conditions that may help to account for performance on the various criterion measures.

With respect to school and college experience, we have included items about major field, academic performance, participation in extra-curricular activities, characteristics of the college environment, the continuity of college experience, the extent of formal education beyond college, aspects of the college experience that stand out in memory, and some corresponding information about the type of high school attended, high school achievement, and participation in various high school extra-curricular activities.

The section dealing with personal information includes the usual kinds of census data such as age, sex, marital status, number of children, race, ethnic background, and occupation. It also includes questions about personal and family background--economic, cultural, political, and religious. In addition, there are items inquiring about the various parts of the country in which the respondents have lived both before and after college and about the extent to which they have traveled in other parts of the world, both before and after college. Finally, in the last part of this section, there is a brief personality test and a brief vocabulary test. The personality test items are related to three characteristics: (a) theoretical orientation, (b) complexity, and (c) autonomy. The vocabulary test calls for definitions of words of various levels of difficulty.

In summary, the total questionnaire with minor modifications, is being administered to samples of incoming freshmen, upperclassmen, and alumni in each of approximately 100 colleges and universities across the country. It includes a broad range of criterion measures, information about the college experience, and data about the background and personal characteristics of the respondents.

Since one of our purposes in this evaluation is to assess institutional effectiveness, we are also obtaining much more data about institutional characteristics than will be obtained from the basic questionnaires. Some of this information comes from studies that have already been made; some of it comes from directories and other sources of demographic information, and some of it will come from instruments which we are still in the process of developing or pretesting. For example, for most of the institutions included in our sample, we have data about the institutional environment obtained from the previous administration of College and University Environmental Scales and from scores obtained by the Environmental Assessment Technique reported in the book, Who Goes Where to College? (Astin, 1965). From directories we are recording a variety of demographic information about the institution such as size, complexity, control, characteristics of faculty, curriculum, and the involvement of the institution in research. These are potentially important aspects of the environment in which learning and development presumably occur.

Beyond these previous and published sources, we are also developing a number of new ways for characterizing the educational environment. At the present time, we can describe six such measures. They are in various stages of development.

The first two are already completed, and they consist of scoring selected items that are included in College and University Environment Scales. One consists of 22 items which, when combined provide an index of Campus Morale. The other consists of 11 items from CUES which, when added together, provide an index called Quality of Teaching and Faculty-Student Relations.

The next two contextual or environmental variables are in the pretesting stage. One of these is concerned with the style of learning or academic effort that is characteristic of a campus. Last year we had a small panel of students from each of four quite different institutions keep a detailed time-log of academic activities for a consecutive seven-day period. In analyzing these time-logs, a number of institutional differences were apparent. In some institutions, for example, there were numerous entries indicating conversations between students and faculty members outside class. At other institutions, there were very few such entries. In some institutions, there were several entries describing conversations with other students about academic matters; whereas, in other institutions, there were very few such entries. Although there were no great differences between institutions in the total amount of time which students reported spending on academic activities, there were noticeable differences in the concentration and intensity of that time--differences, for example, in the frequency of uninterrupted periods of study far into the night and in the amount of study that occurred on week-ends. There were further institutional differences in such activities as the amount of conversation in class, the amount of writing done, and the amount of academically related but non-assigned reading done. These and other items are providing a basis for constructing a brief scale which will identify major differences among the learning styles that characterize each institution.

The other contextual variable that is currently in the pretesting stage is concerned with identifying the institutional stance toward student freedoms and discipline--a stance which ranges from protective and punitive to developmental and permissive.

Two other contextual variables on which we are working are currently in the developmental stage. For one of these, we are trying to devise a way of characterizing the peer group relations and influences that appear to be dominant on the campus. For the other, we are attempting to devise some measure of institutional anonymity and alienation.

As we continue to think about the institutional environment, we shall probably develop still further ways of characterizing it. The point is that much of what has been done in the past has been concerned with relatively small aspects of the environment and educational experience. In the long run, whether one course is taught by lecture or discussion or some combination of the two is probably insignificant. Whether some of the courses make extensive use of television and others do not is probably equally insignificant, as is the question whether the general education requirements must be met by taking a half-dozen specifically designated courses or whether they can be met by selecting a pattern of courses among a larger number of alternatives. It may, on the other hand, be important to know the extent to which students encounter variety rather than sameness in the teaching style of the faculty, and it may be important to know whether the curriculum and other academic regulations are characterized by flexibility or rigidity. What we are looking for are pervasive and stylistic differences that can be felt throughout the college experience rather than particular differences that are evident in only one or a few segments of the college environment.

In the analysis of results from the national field study, we shall be seeking answers to a fundamental but complex question: what personal and background characteristics, what school and college experiences, and what characteristics of institutional environments are associated with each of some 50 or more different and relevant criterion measures? The study we have described differs from previous national studies in two very important ways: first, it includes many, rather than few criterion measures; and second, it includes new kinds of potentially pervasive environmental or contextual measures. The range of criterion measure is really the crucial difference between this and previous studies. It is one thing to identify personal dispositions, educational experiences, and environmental conditions that are predictive of one criterion--such as continuing one's education in graduate school; but it is quite another thing to ask and discover whether or not precisely these dispositions, experiences, and conditions that are positively associated with going to graduate school may be negatively associated with community service, political involvement, or interests in art, music, literature, and drama. It is at least conceivable that the conditions conducive to some kinds of outcomes may be quite different from the conditions conducive to other types of outcomes. It is the concern with different outcomes, each of which is important and relevant to education and the larger society, that gives a new level of complexity to the study we are making.

The essence of evaluation is to place values on, and the important aspect of this statement is that the noun values is plural--not singular. If one's criterion for building a new highway is to locate it so that its construction involves the least amount of earth movement and direct engineering cost, one looks logically at a variety of factors that might be related to this criterion and pays attention to those that have the highest relationship. Other values, such as the displacement of people or the effect on ecology, are ignored because they are not included among the criteria. The result may be a perfectly good research study and, at the same time, a very bad evaluation. The concept of evaluation that underlies our study requires a consideration of many criteria or possible outcomes. If a national study of higher education does not meet these requirements reasonably well, it is not and should not be interpreted as an evaluation.

It is our hope that these studies will make some contribution not only to national perspectives about higher education but also to local institutional self-study and evaluation. The various scales and other measures developed for the study can be used separately or collectively in further studies. Also, since each participating college and university will receive a report of the results obtained from its own students and alumni, local committees or other groups can examine the data and consider their significance. Finally, as we analyze the data from all the colleges and universities, we hope to uncover connections between various practices and outcomes that may have influence on the future effectiveness of institutions.

REFERENCES

Astin, A. W. Who goes where to college? Chicago: Science Research Associates, 1965.

Astin, A. W. Undergraduate achievement and institutional 'excellence.' Science, 1968, 16, 661-668.

Baskin, S. Higher education: Some newer developments. New York: McGraw-Hill, 1965.

Davis, J. A. Undergraduate career decisions. Chicago: Aldine, 1965.

Eckert, R. Outcomes of general education. Minneapolis: University of Minnesota Press, 1943.

Gage, N. L. Handbook of research on teaching. New York: Rand McNally, 1963.

Pace, C. R. Analyses of a national sample of college environments. Final report, USOE Project No. 5-0764. Los Angeles: Graduate School of Education, University of California, 1967.

Reeves, F. W. The University of Chicago survey. Chicago: University of Chicago Press, 1933.

Thistlethwaite, D. L. The effects of college environments on students' decisions to attend graduate school. Final report, USOE Project No. 2993. Nashville, Tenn.: Vanderbilt University, 1968.

Trent, J. W., & Medsker, L. L. Beyond high school. San Francisco: Jossey-Bass, 1968.

Wolfle, D. America's resources of specialized talent. New York: Harper, 1954.