Evaluation, especially when based on a stakeholder-centered model of inquiry, can positively influence institutional change.

Achieving Excellence: How Will We Know?

Clifton F. Conrad, David J. Eagan

How will we know if the curricular changes and innovations we have painstakingly designed, nurtured, and introduced are having any significant effect? Finding out is not a simple task, nor are the results likely to be unambiguous, but such evaluation is of the utmost importance to those committed to bringing about successful change.

Program changes and innovations, whether driven by a private vision or by external mandates, often take on a life of their own once introduced. Partly because curriculum is largely self-propelled and partly because our attention is constantly demanded elsewhere, universities have a tendency to omit the crucial and probably most difficult step in program improvement: assessing impact and consequences. Instead, the assumption is typically made that change, because it is intended to improve, is invariably a good thing. While this may be true to a point, the assumption prevents our finding out exactly what has improved and by how much. It is important that this sourcebook on curricular change include a close look at the process of evaluation as a mechanism for helping institutions to know if excellence has been achieved. This chapter focuses on facilitating evaluation in large institutions, although examples will also be drawn from other settings. We will use the terms *assessment* and *evaluation* interchangeably. Each is sufficiently ambigu-

Conrad, Clifton F. and Eagan, David J. "Achieving Excellence: How Will We Know?" In Improving Undergraduate Education in Large Universities, New Directions in Higher Education, No. 66, edited by Carol H. Pazandak. San Francisco: Jossey-Bass, 1989, pp. 51-63.

ous in meaning to defy precise application, but, in general, we consider assessment, which has lately been tied somewhat narrowly to student outcomes measures, to be a subset of evaluation.

The Challenge of Evaluation

The process of planning for and evaluating change invariably challenges the resourcefulness and stamina of academic institutions buffeted by local, state, and national demands for accountability and curricular improvement. Public institutions in particular have come under intense scrutiny and pressure, and their programs, finances, and missions have been questioned as seldom before. Demands for change and improvement have driven a national scramble to upgrade mission statements and establish new programs, curricula, and, recently, comprehensive assessment initiatives at the program, institutional, and state levels.

Unfortunately, our institutional propensity for launching innovations is only infrequently accompanied by a commitment to assessing the efficacy of what we have done. The University of Minnesota, for example, has a long history of supporting curricular innovations and new programs, but those involved admit that many of those projects have not been accompanied by adequate evaluation. This shortcoming is understandable, given the reward structure and often conflicting expectations of the modern university, but it is clearly not ideal. Innovations and change efforts must be wedded to an evaluation strategy for purposes of determining if a particular project is achieving what it set out to do and of identifying inadvertent outcomes.

Assessment and evaluation are especially nettlesome at large, complex institutions where departmental independence, diffused leadership, and unclear channels of accountability may undermine such monitoring efforts (Dinham, 1988a). Moreover, the pursuit of "excellence" carries so many meanings in today's multiversity that the consensus required in order to make constructive use of evaluation results may be possible only at the least complex levels.

We can best understand the numerous issues embedded in the evaluation of change and innovation by taking into account multiple perspectives on them. Thus, this chapter brings together pertinent literature on evaluation, assessment, program review, and program quality. Relevant examples of evaluation approaches from selected universities are included to ground the discussion in institutional experience.

Choices for Action and Reflection

There are numerous accounts in the literature concerning assessment and evaluation practices in higher education, ranging from statewide

assessment plans (Banta and Fisher, 1984), to institutional assessment plans, to evaluations at the departmental and program level (Conrad and Wilson, 1985). While these project descriptions often contain ideas worth considering in designing evaluations, we are rightly cautioned by Ewell (1985) that "each project is distinctive, and none should prompt direct imitation" (p. 4). Rather, he suggests that we glean an understanding of the choices made in project design and, if possible, the consequences that follow from these choices. In that vein, we propose that a useful approach to evaluation is to examine key decision arenas that deserve attention in any assessment plan.

At large institutions, the strategy for assessing and evaluating change must match the scope of the changes attempted, the institutional context, the people affected, and the financial stake in its success. The larger the project, the greater the challenge in identifying and keeping track of factors that are beyond control. The "realities of the evaluation world," as Patton (1980) calls them, work against the evaluator "who strives to obtain the best possible design and the most useful answers within the real world of politics, people, and methodological prejudice" (p. 18). Obstacles multiply as the scope of a project grows, as those of us in large universities know only too well. One result is that less assessment activity is found at large institutions than at smaller ones (Ory and Parker, 1989).

Notwithstanding the inevitable caveats encountered, a review of the relevant literature presents planners and administrators with a variety of important considerations in the design of an evaluation strategy. In each of these decision arenas, the relative merits of competing philosophies and approaches must be judged. The choices, for example, between qualitative and quantitative measurement and between formative and summative evaluation hold opportunities as well as pitfalls in assessing institutional change.

In this chapter, we discuss six decision arenas where choices made among alternatives are likely to have a significant impact on the evaluation effort. These categories are not exhaustive, nor are they as discrete and clear-cut as the discussion may make them appear. They represent some of the major areas of concern in the literature and professional discourse regarding evaluation, assessment, and program quality. There is a much wider literature that should be explored before final decisions are made; this discussion is intended only to introduce the categories of choice. We bring this "paradigm of choices" (Patton, 1980, p. 20) into consideration both to sharpen awareness of critical issues in evaluation and to lend a measure of concreteness to the discussion through the use of case examples.

Purposes. We assume here that curricular change has been initiated in order to improve existing conditions at an institution and that the findings of an evaluation will be used to further that improvement.

While it can be argued that ultimate uses must await evaluation results, it is nevertheless important to consider the likely uses beforehand in order to aim the study in a direction that can serve the targeted ends. At the same time, an openness to unanticipated uses that arise during or at the conclusion of the evaluation can further enhance improvement.

Building agreement on the purposes of evaluation is a critical preliminary step. The careful consideration of aims and purposes at the outset will expedite subsequent choices as well as guide the overall evaluation design. An evaluation of curricular change can be based on the view that change is iterative, requiring periodic monitoring and corrective input. Or it can focus primarily on the outcomes of the change effort, resulting in a need for summary judgment. These two broad purposes—formative and summative—are not mutually exclusive, but evaluation designers need to choose where to place the strategic emphasis.

Formative. Evaluations that are formative in emphasis (Scriven, 1967) are conducted on an ongoing basis to determine the effectiveness of a project during its implementation and to inform improvements. This approach is favored by those who desire continuous feedback on which to base ongoing adjustments in their implementation strategy. Assessment designed using a formative approach openly reflects a commitment to view evaluation as a kind of learning instrument that guides rather than criticizes. As such, it can serve the interests of those affected by curricular change as well as those responsible for ensuring its effectiveness.

Summative. Evaluating the effectiveness of program change at the end of an arbitrary period constitutes a summative approach (Scriven, 1967). In practice, this is more common than a formative approach and finds wide application in educational settings. Most kinds of student testing and student outcomes measurement address the end product of a course or curriculum, asking, "Did the changes work as intended?" Value-added and norm-referenced measures are typically used to address this question, and input from those being evaluated is often minimal. Decisions about continued funding or program continuance often hinge on the findings of summative judgments.

Involvement and Control. The responsibility for choices made in the design, implementation, and use of an evaluation typically rests with an individual (such as an outside evaluation expert) or group (usually members of the institution) assigned the necessary authority. The decisions that this person or group makes affect all aspects of the evaluation; the interplay of personalities, social and cultural dynamics, and politics needs to be taken into consideration throughout.

In addition to determining where control rests, evaluation designers need to establish the degree of involvement of interested individuals and groups, who can often exert considerable influence. The following alter-

natives illustrate a conceptual division between a broad-based and a more narrow administrative locus of control, although in decentralized institutions like many large universities the distinction is often blurred.

Broad-Based Control. The involvement of stakeholders—defined as any person or group with a direct interest in the outcomes of a change or innovation—has long been recognized as critical to an effective evaluation. Who is more fitting to judge the quality and worth of a program than those most directly involved? Howêver, even minor curricular reform in large universities can affect substantial numbers of people, challenging planners to sample opinions accurately. Because of the inherent difficulties, an honest involvement of stakeholders requires serious effort.

Stakeholders can be incorporated into every step of the process, from the initial identification of purposes to adjustments made during implementation to the translation of findings into policy decisions. We distinguish, of course, between involvement and control, realizing that most of the consequential decisions regarding evaluation design will ultimately be made by those in control. The "responsive" evaluation model (Stake, 1975), which is discussed later, emphasizes this broad-based approach and supports the claim that stakeholder involvement maximizes ownership of both the evaluation and its results.

Administrative Control. The responsibility for virtually all curricular change in colleges and universities resides within an administrative hierarchy. Yet how this responsibility is delegated and the roles that administrators play are open to strategic variation: Strict top-down leadership is at one end of the continuum and dispersed control at the other. Within the context of administrative control, an evaluation can be designed to encourage either broad or narrow participation among the units that are directly affected.

Strong administrative leadership has been cited as a key factor influencing success in program implementation and evaluation (Conrad, 1978). The success of outcomes assessment initiatives under way in a number of states has been attributed to the singular commitment of a university president or other high-ranking individual (Provost's Steering Committee on Assessment, 1987). Administrative leaders can offer a vision of the potential accomplishments of program innovation and change and can have significant impact on involvement and long-term success.

Focus and Scope. Inevitable limits in resources and time make it advisable to restrict the scope of a project to those areas of greatest concern that can be dealt with fully. These constitute the "short list" of areas to be examined. A short list helps ensure that the evaluation will concentrate on areas where it can have the greatest effect.

Among the factors influencing project scope, institutional politics

will probably play a significant role. Accordingly, assessing the political climate may be necessary before deciding what aspects of change to evaluate. At large universities in particular, there will often be political expedients—that is, larger payoffs—linked to certain areas targeted for evaluation. But, if such choices are made only for political reasons, they may result in what Edelman (1977) terms "words that succeed and policies that fail." Within the context of these concerns, the scope of an evaluation will range from broad to narrow.

Comprehensive. An evaluation of curriculum change can take a broad, comprehensive view, addressing a host of factors that affect the implementation and outcomes of change. As an example, the assessment plan at the University of Arizona places major emphasis on the total undergraduate experience, taking into account the initial capabilities of students, the academic curriculum and extracurriculum, elements of the institutional environment, and a broad range of student outcomes (Conrad, 1987). The newly established Center for Research on Undergraduate Education coordinates these efforts. At Arizona, where the aim of assessment is institutional improvement, a comprehensive approach is deemed essential in revealing overall context and ongoing change. Such large projects are characterized by greater complexity and ambiguity, but they also hold the potential for richer analysis.

Specific. An evaluation effort may be aimed specifically at measuring and judging the effectiveness of curricular change from as narrow a perspective as desired. The assessment program at the University of Arizona, in addition to its more global perspective, provides for specific evaluation efforts to respond to "current questions" that arise within the institution (Dinham, 1988b); shorter-term, focused attention is given to special problems and concerns. Evaluations designed to judge the effectiveness of specific programs or isolated curricular modifications may be less hampered by complexity of the sort found in comprehensive efforts, but their findings may also have more limited application.

Evaluation Models. Criteria, evidence, and judgment are the common threads uniting all evaluation efforts into a sort of kinship, however distant the relationship may appear. They are at the core of all evaluations, regardless of the rhetoric and packaging that can make different approaches appear to be polar opposites. This is not to imply that the methods chosen to pursue these threads make little difference. On the contrary, hundreds of research articles and books are devoted to the exploration of the many concerns and debates associated with the evaluation task. Although much of the earlier literature centered on explicating methodological approaches to evaluation, current practice and theory often explore evaluation from the context of utilization and decision making (Patton, 1978; Weiss, 1988; Shapiro, 1986).

Complex projects at large institutions may call for several concurrent evaluative strategies, involving the use of methodologies and orientations from two or more models. The selection of models needs to be based on their usefulness in addressing the project emphases generated from previous design choices (on purpose, scope, focus, and stakeholder involvement). The many overviews of evaluation models in the literature can provide a helpful orientation to the field and can aid in determining which approach is best suited for a particular institution (Gardner, 1977; Madaus, Scriven, and Stufflebeam, 1983; Conrad and Wilson, 1985; Shapiro, 1986). The following is a summary of four major evaluation models, based on the typology of Conrad and Wilson (1985).

Goal-Based Model. Variations on this model predominate in evaluation efforts in higher education and are grounded in the work of Tyler (1949). Evaluating the attainment of objectives is at the heart of the goal-based approach. Previously established program goals, objectives, and standards of performance are identified, program outcomes are measured, and a judgment is made based on the congruence or discrepancy between planned objectives and demonstrable outcomes (Gardner, 1977). While a goal-based approach is typically summative in intent, Provus (1971) has expanded the scope of the model to serve formative purposes as well by including analysis and interpretation of intended program processes in addition to outcomes.

Responsive Model. The responsive approach is organized broadly around "the concerns and issues of stakeholding audiences" (Guba and Lincoln, 1981, p. 23). Originally developed by Stake (1975), the responsive model stresses that evaluation efforts should not be driven narrowly by program goals but rather that an understanding of "unintended effects" (Scriven, 1973) and of stakeholder concerns is necessary for interpreting outcomes. The design of a responsive evaluation is an ongoing process, since each step is informed in part by previous activity (Guba and Lincoln, 1981).

Decision-Making Model. Some educators believe that the ultimate purpose of an evaluation effort is to inform administrative decisions. Thus, this model is organized around the decision-making process. The most widely known decision-making model is the Context, Input, Process, Product (CIPP) model (Stufflebeam and others, 1971), which holds that the different types of decisions inherent in the evaluative process require different kinds of evaluation activities. Four types have been identified: context evaluation, which assists decision makers in determining goals and objectives; input evaluation, which helps clarify alternative ways of achieving program goals and objectives; process evaluation, which provides feedback to decision makers; and product evaluation, which provides decision makers with information as to whether a program should

be continued, modified, or terminated. While the CIPP approach has not been widely applied in higher education, the number of institutions initiating decision-oriented evaluation is increasing (Conrad and Wilson, 1985).

Connoisseurship Model. In many instances, evaluations are entrusted to persons whose expertise qualifies them to judge the relative merits of a program in all its complexity, subtlety, and nuance. Under the connoisseurship model (Eisner, 1976), the connoisseur alone guides the evaluation, balancing and comparing information gleaned through documents, interviews, and observation with a continuous, more intuitive awareness and sense of appreciation, which Eisner likens to the appreciation of art. The visits of accreditation review teams are based partly on this model; team members' extensive experience gives them a connoisseurship on which to base their judgments about program quality.

Evaluation Emphasis. An important arena of choice in evaluation concerns the emphasis on program versus that on product—in other words, whether to assess curricular change from the viewpoint of the program itself or from its outcomes. Each orientation provides different kinds of information about quality and effectiveness that can serve decisions from either a formative or summative perspective. Both emphases can be combined in an evaluation, multiplying the complexity of the task but providing a more complete analysis.

The "assessment movement" in higher education, with its expanding base of scholarly literature and growing institutional practice, illustrates how this difference in evaluation emphasis plays out in institutional settings (Ewell, 1987). While most institutions have focused their attention on the outcomes of the undergraduate experience, many have attempted to look as well at the learning environment and other conditions that may influence outcomes. The choosing and implementing of an assessment emphasis or "model" is no simple matter, however, as Ewell and Boyer (1988) imply in their examination of recent assessment experiences in five states (Colorado, Missouri, New Jersey, South Dakota, and Virginia).

Academic Program, Environment, and Student Characteristics. The quality and impact of the academic and institutional environment, which can have significant effects on student achievement, are troublesome to measure. Because of this difficulty and also because of mandates for assessments and evaluations that emphasize data on outcomes, little systematic effort has targeted this area.

The comprehensive assessment project at the University of Arizona, however, provides an instructive example. Through a variety of measurement strategies, the Arizona Plan (Conrad, 1987) seeks to identify and assess factors in three broad areas that influence student outcomes: student characteristics, including knowledge, attitudes, and intellectual

skills; components of the undergraduate experience, specifically general education, the major, and the extracurriculum; and institutional environment, composed of climate, resources for learning, faculty, students, and curriculum. Examining these factors provides a context within which decision makers can interpret evaluation results.

Outcomes. In higher education, the prevailing emphasis in evaluation studies is on measuring student outcomes. Most efforts seek to assess educational proficiency or gain in student knowledge and skills, though some strive to document student growth in other areas as well. Alverno College is frequently cited for its commitment to a comprehensive student assessment strategy that is integrated fully with the curriculum and that measures such student abilities as problem solving, valuing, and taking environmental responsibility (Alverno College Faculty, 1985; Mentkowski and Loacker, 1985). Similarly, the "talent development" approach to achieving educational excellence (Astin, 1985; Jacobi, Astin, and Ayala, 1987) takes a holistic approach to measuring long-term student growth and development.

Most outcomes assessment approaches, however, place primary emphasis on testing and other quantifiable measures of student achievement. The value-added method, which assesses gain in knowledge and skills over time as measured using pretests and posttests, forms the core of the assessment program at Northeast Missouri State University (McClain and Krueger, 1985). At the University of Tennessee at Knoxville, a performance-based funding assessment effort uses value-added tests such as the American College Testing College Outcomes Measures Project (ACT COMP) to measure student achievement in both the major and general education (Banta, 1985). In many institutions, nationally normed tests such as the GRE and MCAT are also employed. In an instructive treatment of this topic, Jacobi, Astin, and Ayala (1987) address many of the limitations of outcomes measurement and urge the choosing of testing instruments that match the goals and values of an institution.

Methods of Measurement. Effective evaluation requires that measurements be as accurate and meaningful as possible. The literature in evaluation and related fields, especially in critical reviews and meta-analyses, contains extensive debate over methodology, much of it centered on problems of measurement. An ongoing quarrel focuses on the tendency in education, as in the social sciences overall, to value quantitative over qualitative techniques (Cook and Reichardt, 1979; Patton, 1980; Bednarz, 1985). The design of an evaluation needs to address the balance between these two techniques, weighing the relative merits of each.

Quantitative. Quantitative measurement is prevalent in evaluation and assessment in which data are based on test scores, survey statistics, Likert-type scales, and other numerically oriented measures. Such data

are frequently viewed as the most scientific and, hence, the most valid and reliable kinds of information. In goal-based evaluations, for example, many of the criteria used to compare objectives with outcomes are grounded in some quantifiable measure of performance. An overarching concern must be that such quantitative data reflect reality as closely as possible.

Qualitative. Increased attention has been focused in recent years on naturalistic and qualitative approaches to evaluation (Lincoln and Guba, 1985). Interviews, open-ended survey questions, and participant observation are qualitative methods that can add the depth and meaning that is not easily captured through quantitative techniques. Moreover, case studies (Merriam, 1988) and ethnographic evaluation (Tierney, 1985) are qualitative approaches aimed at extracting ideographic and cultural meaning from educational settings.

Conclusion

The foregoing arenas for choice are perhaps best conceived as overlapping pieces of a larger picture, each existing in a dynamic balance with the others. Making choices among them is not necessarily an either/or matter but should be viewed as a process of balancing alternatives that seem most fitting for a given set of circumstances.

Historically, many of us in higher education have been more adroit at initiating change—or, rather, forging ahead with new ideas—than we have been at evaluating realistically and thoroughly the changes we have introduced. We are, first, creators and planners and, second, evaluators and formal critics. In higher education today, there is probably one evaluation committee for every dozen committees planning for change.

In many ways, implementing high-quality programs is a never-ending, Sisyphean task. Just when we think excellence is within reach, new priorities and circumstances often arise that refigure our institutional trajectory. Evaluation can take away some of the guesswork in implementing our ideas and can help guide us into more promising directions for future change and innovation. Furthermore, deliberate assessment and evaluation can reveal aspects of quality and areas of concern that were unanticipated in the original project design. Jacobi, Astin, and Ayala (1987), for example, describe several unanticipated effects of value-added assessment, such as increased student test anxiety, awareness of intellectual development, and better test-taking skills. From the perspective of these researchers, successful evaluation projects not only measure effectiveness but also produce an impact themselves. Similarly, we believe that carefully designed evaluation efforts can have a positive influence on overall institutional development, as well as on the specific change they originally set out to assess.

In closing, we urge that serious attention be directed to the evaluation of curricular change in order both to understand the consequences of that change and to recognize the significance of change to all relevant constituent groups. Such attention will lead, we hope, to a model of inquiry centered on the stakeholder (Conrad, 1989), which in turn will lead to a broader consensus about and commitment to what it takes to achieve excellence.

References

Alverno College Faculty. *Assessment at Alverno College.* Milwaukee, Wisc.: Alverno Productions, 1985.

Astin, A. W. *Achieving Educational Excellence: A Critical Assessment of Priorities and Practices in Higher Education.* San Francisco: Jossey-Bass, 1985.

Banta, T. W. "Use of Outcomes Information at the University of Tennessee at Knoxville." In P. T. Ewell (ed.), *Assessing Educational Outcomes.* New Directions for Institutional Research, no. 47. San Francisco: Jossey-Bass, 1985.

Banta, T. W., and Fisher, H. S. "Performance Funding: Tennessee's Experiment." In J. K. Folger (ed.), *Financial Incentives for Academic Quality.* New Directions for Higher Education, no. 48. San Francisco: Jossey-Bass, 1984.

Bednarz, D. "Quality and Quantity in Evaluation Research: A Divergent View." *Evaluation and Program Planning,* 1985, *8* (4), 289–306.

Conrad, C. F. "A Grounded Theory of Academic Change." *Sociology of Education,* 1978, *51* (2), 101–112.

Conrad, C. F. *Plan for Assessing Undergraduate Education at the University of Arizona.* Tucson: Task Force on Assessment of the Quality and Outcome of Undergraduate Education, University of Arizona, 1987.

Conrad, C. F. "Meditations on the Ideology of Inquiry in Higher Education: Exposition, Critique, and Conjecture." *Review of Higher Education,* 1989, *12* (3), 199–220.

Conrad, C. F., and Wilson, R. F. *Academic Program Reviews.* ASHE-ERIC Higher Education Report, no. 5. Washington, D.C.: Association for the Study of Higher Education, 1985.

Cook, T. D., and Reichardt, C. S. *Qualitative and Quantitative Methods in Evaluation Research.* Beverly Hills, Calif.: Sage, 1979.

Dinham, S. M. "Issues and Methodological Concerns for Undergraduate Assessment in Large Research Universities." Paper presented at the annual meeting of the Association for the Study of Higher Education, St. Louis, November 3–6, 1988a.

Dinham, S. M. *Summary of Assessment Activities at the University of Arizona.* Internal Report, no. 4. Tucson: Center for Research on Undergraduate Education, University of Arizona, 1988b.

Edelman, M. *Political Language: Words That Succeed and Policies That Fail.* Orlando, Fla.: Academic Press, 1977.

Eisner, E. W. "Educational Connoisseurship and Criticism: Their Form and Function in Educational Evaluation." *Journal of Aesthetic Education,* 1976, *10* (3–4), 135–150.

Ewell, P. T. (ed.). *Assessing Educational Outcomes.* New Directions for Institutional Research, no. 47. San Francisco: Jossey-Bass, 1985.

Ewell, P. T. "Assessment: Where Are We?" *Change,* 1987, *19* (1), 23–28.

Ewell, P. T., and Boyer, C. M. "Acting Out State-Mandated Assessment: Evidence from Five States." *Change,* 1988, *20* (4), 40–47.

Gardner, D. E. "Five Evaluation Frameworks: Implications for Decision Making in Higher Education." *Journal of Higher Education,* 1977, *48* (5), 571–593.

Guba, E. G., and Lincoln, Y. S. *Effective Evaluation: Improving the Usefulness of Evaluation Results Through Responsive and Naturalistic Approaches.* San Francisco: Jossey-Bass, 1981.

Jâcobi, M., Astin, A. W., and Ayala, F., Jr. *College Student Outcomes Assessment.* ASHE-ERIC Higher Education Reports, no. 7. Washington, D.C.: Association for the Study of Higher Education, 1987.

Lincoln, Y. S., and Guba, E. G. *Naturalistic Inquiry.* Beverly Hills, Calif.: Sage, 1985.

McClain, C. J., and Krueger, D. W. "Using Outcomes Assessment: A Case Study in Institutional Change." In P. T. Ewell (ed.), *Assessing Educational Outcomes.* New Directions for Institutional Research, no. 47. San Francisco: Jossey-Bass, 1985.

Madaus, G. F., Scriven, M., and Stufflebeam, D. L. (eds.). *Evaluation Models: Viewpoints on Educational and Human Services Evaluation.* Boston, Mass.: Kluwer-Nijhoff, 1983.

Mentkowski, M., and Loacker, G. "Assessing and Validating the Outcomes of College." In P. T. Ewell (ed.), *Assessing Educational Outcomes.* New Directions for Institutional Research, no. 47. San Francisco: Jossey-Bass, 1985.

Merriam, S. B. *Case Study Research in Education: A Qualitative Approach.* San Francisco: Jossey-Bass, 1988.

Ory, J. C., and Parker, S. "Survey of Assessment Activities at Large Research Universities." Paper presented at the annual meeting of the American Educational Research Association, San Francisco, March 27–31, 1989.

Patton, M. Q. *Utilization-Focused Evaluation.* Beverly Hills, Calif.: Sage, 1978.

Patton, M. Q. *Qualitative Evaluation Methods.* Beverly Hills, Calif.: Sage, 1980.

Provost's Steering Committee on Assessment. *Assessment of Undergraduate Student Outcomes at UMC.* Columbia: University of Missouri, 1987.

Provus, M. M. *Discrepancy Evaluation.* Berkeley, Calif.: McCutchan, 1971.

Scriven, M. S. *The Methodologies of Evaluation.* AERA Monograph Series in Curriculum Evaluation, no. 1. Chicago: Rand McNally, 1967.

Scriven, M. S. "Goal-Free Evaluation." In E. R. House (ed.), *School Evaluation: The Politics and Process.* Berkeley, Calif.: McCutchan, 1973.

Shapiro, J. Z. "Evaluation Research and Educational Decision Making: A Review of the Literature." In J. C. Smart (ed.), *Higher Education: Handbook of Theory and Research.* Vol. 2. New York: Agathon Press, 1986.

Stake, R. E. (ed.). *Evaluating the Arts in Education: A Responsive Approach.* Columbus, Ohio: Merrill, 1975.

Stufflebeam, D., and others. *Educational Evaluation and Decision Making.* Itasca, Ill.: Peacock, 1971.

Tierney, W. G. "Ethnography: An Alternative Evaluation Methodology." *The Review of Higher Education,* 1985, *8* (2), 93–106.

Tyler, R. W. *Basic Principles of Curriculum and Instruction: Syllabus for Education 360.* Chicago: University of Chicago Press, 1949.

Weiss, C. H. "Evaluation for Decisions: Is Anybody There? Does Anybody Care?" *Evaluation Practice,* 1988, *9,* 5–19.

Clifton F. Conrad is professor of higher education at the University of Wisconsin-Madison.

David J. Eagan is a doctoral student in higher education, University of Wisconsin-Madison.