University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

POD Quarterly: The Journal of the Professional and Organizational Development Network in Higher Education

Professional and Organizational Development Network in Higher Education

7-1-1979

The Practice of Evaluation

Glenn F. Nyre

Follow this and additional works at: http://digitalcommons.unl.edu/podqtrly

Part of the Higher Education Administration Commons

Nyre, Glenn F., "The Practice of Evaluation" (1979). *POD Quarterly: The Journal of the Professional and Organizational Development Network in Higher Education.* Paper 13.
http://digitalcommons.unl.edu/podqtrly/13

This Article is brought to you for free and open access by the Professional and Organizational Development Network in Higher Education at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in POD Quarterly: The Journal of the Professional and Organizational Development Network in Higher Education by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Evaluation

The Practice of Evaluation

GLENN F. NYRE

I was over thirty before I finalized my decision not to become a minister. But even now, as some of you know, I still occasionally give in to the urge to preach. Thanks to an invitation from the editor of the *POD Quarterly,* I now have an opportunity to do so on a regular basis—and on a topic close to my heart.

Evaluation is part and parcel of our POD jargon. Unfortunately, as is the case with some other concepts and practices we borrowed from our predecessors and colleagues in other fields, it is extremely misunderstood. Many of us have become "overnight experts" in and about things we had no knowledge of when "the movement" began, and evaluation is no exception. In fact, I know of only three people in POD that have had any formal training in evaluation. Admittedly, it is my bias that evaluation is a more complex and (dare I say it?) more important skill than some others currently being plied in the trade. Thus my willingness to write this column.

There have been many articles and conference presentations about the evaluation of professional development activities in the past few years, but they have typically been one of three types: 1) diatribes about the lack of evaluation in the profession, 2) simplistic, "experimental" studies comparing "X" characteristics of teacher evaluation forms, or 3) jargon-confounded articles written by evaluators to impress other evaluators. As a result, the art of evaluation has not been advanced among professional development practitioners.

I admit to having been a contributor to the proliferation of these articles in the past, and in fact, have even gone so far as to combine all three types in one article. Hopefully, my retribution will be made through this column. I do *not* intend to write "how to evaluate a

professional development program" articles. We have been that short-circuit route, and it has not served us well because there are no universal methods for that purpose. As with most other things in life, the procedure which is best for your program depends on a lot of concerns and constraints relative to your situation.

Rather, I will, in subsequent editions, summarize the major evaluation theories, as well as discuss the advantages and disadvantages of the various methods and models in vogue today and on the horizon. With this background, you may be able to select evaluation strategies appropriate to your program, and possibly even come up with some hybrid procedures of your own.

The content of the columns will be sequential and develop in the reader a cumulative knowledge of the field. This is not only a more-than-acceptable theory of learning, but will also serve as a strategy to get people who begin their subscriptions late to order back issues. This introductory column will present a brief historical overview of the field and define certain terms which will be used in subsequent columns.

Historical Overview

Formal evaluation has a very long history, dating back to 2000 B.C. when Chinese officials began administering civil service examinations. Although it remains undocumented, Clare Rose has traced the beginnings of evaluation back to Moses, suggesting that he must have carefully evaluated the consequences of risking the perils of foreign travel at such an advanced age. The history of evaluation in this country is more germane to this column, however, and can be traced through four landmark dates.

The first formal educational evaluation in the United States was conducted by Joseph Mayer Rice in 1887. He developed a spelling test and administered it to 30,000 students in an attempt to show that student achievement was not related to the amount of time spent in spelling drills. His name will never be in your children's history books, but I like to promote him because he was not an educator, but rather, a pediatrician who was fed up enough with the educators of his day to finance his own study.

Unfortunately, the practice of evaluation was not greatly furthered as a result of Rice's contribution, since all that happened for the next forty years was the institutionalization of standardized tests

for school children. The field did not really advance until the 1930s when Ralph Tyler promoted an approach to evaluation which went beyond giving tests.

Tyler conceived of evaluation as the process of determining the degree to which the goals of a program have been achieved. Although the rest of the world was still not aware of the glories of evaluation and evaluators, a sort of "underground" profession began to develop—somewhat akin to the early POD movement, except that evaluation remained in this state for about another 35 years. It was not until after the launching of Sputnik that evaluation was recognized as a semi-worthy activity. After the passage of the Elementary and Secondary Education Act (ESEA) of 1965, evaluation began to flourish because of the clause contained in the thousands of grants awarded which required a formal evaluation as a condition of the award.

Needless to say, so-called evaluator$ came out of the woodwork to cash in on this event. Unfortunately, there was at this time no such thing as a professional evaluator (hold the applause, please); academics trained in research or measurement were drafted to conduct the evaluations. Their primary products (measurement-filled reports) still sit on shelves today, and their other products (graduates) still sit on the same designs.

But into this great void came people like Campbell and Stanley, Jim Popham, Michael Scriven and Carol Weiss. These people are among the leaders in educational evaluation today. I would also mention Guba, Stufflebeam and Airasian, but you might think I am making up names. I would also mention myself, but you might think I am egotistical. Regardless, we all find ourselves in the midst of a booming profession which is full of confusion, conflict and controversy. We do not share a common philosophy, focus or even terminology. But we do all share one characteristic—a love of building models. The psychological needs of some of us may have been arrested in the air-place and car-building stages of our youth; others to us are simply die-hard Rube Goldberg fans who are intrigued by mazes of convoluted lines, arrows and dots.

Evaluation models are as prolific as rabbits, and they procreate about as speedily. No longer do people develop an idea or test an approach. Instead, they develop a model. Often spawned from combinations of several other models, some from other disciplines, they

become progressively more grandiose in their complexity, more esoteric in their terminology and more pompous in their names. Many of these so-called models, of course, are not really models, but rather, descriptions of processes or approaches to program evaluation.

The array of evaluation models from which we may choose would, if nothing else, provide a marvelous tongue-twisting party game. Just imagine what it would sound like if someone who'd had too much to drink were to chant in mantra form the names of evaluation models and approaches. We have democratic evaluation, responsive evaluation, transactional evaluation, modus-operandi evaluation, holistic evaluation, discrepancy evaluation, goal-free evaluation, and adversary evaluation. There is the Countenance Model, the Differential Evaluation Model, the Priority Decision Model, the Trade-Off and Comparative Cost Model, the Systems Approach Model, the Apportionment Model, and the Cost Utility Model. There are Ontological Models, Synergistic Models, and Ethnographic Models. And this is only a partial list.

Nobody said it was going to be easy. But fear not. These columns shall lead you out of the depths of darkness into a never-never land heretofore known only to evaluators and other perpetrators of white collar crime. If you have read this far, there is no turning back. You will never be satisfied until you know the full Gospel of Evaluation according to Nyre. And in order to begin preparing for the meatier issues of the next column, I would suggest that you acquaint yourselves with the terms discussed below.

Terminology

Definitions and distinctions are not idle concerns, I assure you. Misunderstandings of these words and phrases are at the heart of many unnecessarily heated debates in the profession. Even the most basic terms, such as measurement, assessment, accountability and even evaluation itself are used interchangeably and often incorrectly. Is it any wonder that in some quarters evaluation has not yet been recognized as a legitimate enterprise?

Accountability. Accountability is concerned with furthering the educational effectiveness of school systems. My dictionary shows the synonym of accountability to be "responsibility." Educational accountability thus represents the educators' acceptance of respon-

sibility for the consequences of the educational system entrusted to them by the public. Evaluation is an intrinsic part of accountability. Program effectiveness must be evaluated to provide information for teachers, administrators and program directors, as well as legislators and other officials who allocate the funds for the programs and for the public who provides the funds through their tax dollars. Accountability is usually a condition *requiring* evaluation; but accountability is not equivalent to evaluation.

Measurement. Measurement is often equated with evaluation, since so many of the early evaluation reports consisted primarily of measurement data. But measurement is static—it is the act or process of determining the extent, dimensions, quantity, or capacity of something at one point in time. In education, measurement is the act of determining the extent to which an individual has learned or the degree to which an individual possesses a certain characteristic, ability, or talent. Measurement is usually part of the evaluation process, providing useful data for evaluation, but again, the two terms are not equivalent.

Assessment. Like measurement, the term assessment is often used interchangeably with evaluation, and several major evaluation projects have been referred to as "National Assessments." Assessment is really more akin to measurement, however, and refers to the process of gathering and collating the data. Assessment has a narrower meaning than evaluation and a broader meaning than measurement. In addition to the act of measurement, assessment involves the qualitative judgment of determining what and how to measure as well as the process of putting the data into an interpretable form.

Evaluation. Everyone knows what evaluation is. Or do they? There are several definitions of evaluation, and the one to which evaluators subscribe affects the way in which they approach and carry out their evaluations. The various definitions also provide conceptual bases for the different models of evaluation. Although there are still a few educators who subscribe to the measurement definition (e.g., Thorndike, Ebel) most model builders and evaluation writers cluster around three major definitions: 1) those that define evaluation as an assessment of the discrepancy between objectives and performance (Metfessel and Michael; Provus; Stake; Tyler); 2) those that focus on outcomes and define evaluation as an

assessment of outcomes, intended or otherwise (Popham; Rose; Scriven); and 3) those who are decision oriented defining evaluation as the process of obtaining and providing information for decision makers (Alkin; Cronback; Guba and Stufflebeam).

If these names seem strange to you (other than the fact that the field has attracted so many people with funny names), do not worry. Each of these "schools" of evaluation thought and the writings and models of their proponents will be discussed in subsequent columns, so you will never have to feel inadequate in the presence of a real-life evaluator again—at least as far as basic knowledge is concerned.