

ways of achieving widespread implementation and, as a general policy, are to be preferred to the negative approach that emphasizes enforcement and sanctions. Nevertheless, educational evaluators should also recognize that if they do not effectively regulate their activities, other groups, especially courts and governmental bodies, will. In general, those concerned with advancing the practice of educational evaluation should carefully consider how they can promote and ensure the implementation of the *Standards*.

In summary, standards are a vital aspect of a secure profession. Over the last few years, a variety of organizations concerned with improving the professional practice of educational evaluation have supported an extensive effort to develop a first set of standards for this field. These standards are available and should immediately prove useful. On the other hand, the setting of standards must be an ongoing process if such guidelines are to promote and not stifle our field. Also those involved in the educational evaluation profession should take steps to ensure that the *Standards* will be effectively implemented. I have described the Joint Committee's plan for a sustained effort and have noted that ENet will be represented on the reconstituted Joint Committee.

I appreciate Barbara Davis' efforts to secure the reactions contained in this section of *EN*. I am sure that the Joint Committee will find them instructive regarding how the present standards can be improved.

## **The *Standards*: A General Review**

**Lee Sechrest**

*Director, Center for Research on Utilization  
of Scientific Knowledge  
Institute for Social Research  
University of Michigan  
Ann Arbor, MI 48104*

The publication of standards for evaluation may eventually come to be regarded as a step of considerable importance in the development of program evaluation as a professional activity. Undoubtedly, fault will

be found with the standards, and surely over time they will be amended and in various ways improved. Their very publication does, however, give program evaluators and those who commission and use evaluations a benchmark against which to judge their activities and products. The *Standards* are likely to prove of considerable value in teaching and training of evaluators, and they will be likely to influence the content of courses and training programs in a variety of ways. Certainly they will sensitize all of those involved in evaluations to an extraordinarily wide range of issues and problems that must at least be considered, if not resolved, if evaluation is to become a dependably worthwhile, productive activity.

These standards are the product of what was clearly a thoughtful, carefully delineated approach to the problem by a knowledgeable group of experts who took their charge quite seriously. The consistent format of the *Standards* makes them especially easy to use and will also make them easy to amend or elaborate upon as the occasion arises. Individual evaluators or teachers will find it easy to add their own caveats or pitfalls or to supplement illustrations with personal experience. The *Standards* are, at least taken individually, highly realistic and commanding of adherence. The illustrations are interesting in themselves and ring true. Virtually anyone with some experience in evaluation will have encountered instances quite similar to the cases that are described in the *Standards*. Evaluators will, in fact, have encountered instances of one sort or another that are worse than those pointed to as bad examples.

Many of the problems that arise in evaluation are or should be avoidable. The *Standards* might also be of considerable help in negotiating evaluations with those wishing to contract for them. Those who commission or pay for evaluations ought to know ahead of time the problems that may arise and have a basis on which to negotiate such difficult matters as political controversies, identification of constituents, and ways of eliciting cooperation of staff. It should help evaluators considerably to have a concise, well-written set of standards that can be displayed to those with whom they contract to provide services.

Standards are essentially moral prescriptions and like any set of commandments they are easier to state than to follow. The standards proposed here are demanding and few will be able to live up to all of them. It is to be hoped, however, that unlike other sets of commandments, these will at least inspire serious attempts to live up to them. The difficulty with the *Standards* probably begins with their sheer number. Judeo-Christian morality could be expressed in ten commandments. In evaluation we require thirty standards. As if the number of standards were not enough, they are interrelated in complex ways so that some of

them on some occasions may actually be incompatible with each other, a fact which the Joint Committee which prepared the standards admits quite freely. For example, the standard urging full and frank disclosure of outcomes of evaluations may run afoul of the standard that requires that political and other contextual realities be taken into account in publishing the results of evaluations. Moreover, as reasonable as the standards appear to be when examined in the abstract, some of them may be more than easy to overlook in the hurly burly of actual evaluation planning and implementation. At least to this reviewer, the illustrations themselves forewarn of some of the problems. Although the analyses of each of the illustrative cases are sensible and persuasive, a little reflection will suggest to many experienced evaluators that hindsight is a much better guide to action than foresight. For example, several of the standards relate in one way or the other to the identification of appropriate participants in and audiences for evaluations. The standards all urge that appropriate participants and audiences be identified so that their interests and needs can be taken into account. Some of the illustrations of problems indicate that it may be possible for even a well-meaning and experienced evaluator to be surprised at just who does have a stake in an aspect of an evaluation once the evaluation is underway or once a report has been published. Moreover, personal experience indicates that those contracting for an evaluation may want, sometimes legitimately, to ignore or to exclude a particular potential audience. About all an evaluator may be able to do in some circumstances is to warn of the troubles that lie ahead. Nonetheless, despite these complaints, this reviewer regards the *Standards* as provocative and themselves a forewarning to all those involved in evaluation of the problems that lie ahead if an evaluation is not carefully thought through, executed, and reported.

This reviewer is personally disappointed that the *Standards* do not contain specific references to the appropriate experimental (or quasi-experimental) designs for evaluation. Two of the standards pertain to the requirement for using reliable and valid measures. There is no standard that calls for the use of only valid and dependable designs. A standard that called for use of the best possible research design would have been welcome. Related to that is, again in the view of this reviewer, the need for a standard to require an evaluator to explain fully to sponsors and other audiences the exact nature of the conclusions likely to be permitted by the evaluation as planned and designed. For a variety of reasons, planned evaluations may fall short of ideal and can be counted on to produce less than firm conclusions. A good evaluator may

accept in advance that those responsible for programs are contracting for evaluations likely to lead to murky conclusions. They do need to be forewarned, however, what it is that they are contracting for.

One hopes that, unlike the ten commandments, these standards will not be graven in stone and that they therefore will be subject to revision, amendment, and change of an evolutionary nature as the field itself grows and develops. If that is the case, then this document will have fulfilled its purpose and in the long run will be recognized as a sign of the coming of age of evaluation as a scientific and professional activity. Given that these standards were prepared by a committee and not pronounced by a single voice from a burning bush, they are an excellent beginning.

## Setting Standards for Educational Evaluators

**Robert E. Stake**

*Director, Center for Instructional Research  
and Curriculum Evaluation*

*University of Illinois  
270 Education Building  
Urbana, IL 61801*

I have four questions about setting standards for the conduct of educational evaluation:

(1) When should the profession attempt to control the individual evaluator?

(2) Should standards be set even in the absence of means of measuring accurately and even when suitable sanctions against offenders do not exist?

(3) How uniform should standards be?

(4) How prespecified should standards be?

I will respond to each of these questions, then comment on the *Standards* effort, and finally mention an obvious alternative to a formal compedia of standards.