DOCUMENT RESUME

ED 183 574 TM 010 072

AUTHOR Lincoln, Yvonna S.; Guba, Egon G.
TITLE The Distinction Between Merit and Worth in Evaluation.
PUB DATE Sep 79
NOTE 21p.; Paper presented at the Annual Meeting of the Evaluation Network (5th, Cincinnati, OH, September 24-26, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Conceptual Schemes; Cultural Pluralism; *Evaluation; Evaluation Criteria; Evaluation Needs; Evaluators; Formative Evaluation; Summative Evaluation; *Values
IDENTIFIERS *Merit; *Worth

ABSTRACT

Valuing in evaluation encompasses two distinct senses of the word, denoted by the terms merit and worth. Merit may be defined as an entity's inherent, intrinsic, context-free value, while an entity's worth is defined as its contextually determined, place-bound value. Determining an entity's merit may take place whenever a number of experts are assembled. Worth can only be determined by viewing the entity in operation or on site. Thus, while merit may be determined in any number of ways, worth can be determined only by intensive field studies on site. And field studies often call for naturalistic, not scientific, approaches. Although it would seem that merit and worth are identical to formative and summative dimensions, they are orthogonal. It is therefore possible to create a 2 X 2 table and generate four distinct types of evaluation: formative merit evaluation, formative worth evaluation, summative merit evaluation, and summative worth evaluation. Each of the four types of evaluation serves distinctly different purposes and is addressed to different audiences and stakeholders. (Author/BW)

THE DISTINCTION BETWEEN MERIT AND WORTH IN EVALUATION

Yvonna S. Lincoln

School of Education
University of Kansas

and

Egon G. Guba
School of Education
Indiana University

A paper presented at the Annual Meeting of the Evaluation Network
Cincinnati, Ohio, September 1979

Some Definitions

The root of the term evaluate suggests that the function of evaluation is to place a value on the thing being evaluated. But there are two senses in which an entity's value may be conceptualized. On the one hand, the object may have value of its own, implicit, inherent, independent of any requirements of applicability or use. So for example, "pure" science, as the adjective pure suggests, has value untainted by any considerations outside the sphere of science itself. It is enough that scientists appreciate (esteem, admire) the pure discovery or development for its own sake. We shall apply the term merit to this kind of intrinsic, context-free value.

On the other hand, the entity may also have value within some context of use or application; thus, "applied" science is undertaken for the sake of solving some practical problem. The products of applied science are valued to the extent to which they provide solutions, i.e., have utility in a practical context. While some scientists may denigrate these products as having "no theoretical significant" (i.e., as not being meritorious), they may nevertheless have great value in an engineering or indutrial setting. We shall apply the term worth to this kind of extrinsic or context determined value.[1]

Some additional examples will help to make the distinction we are drawing clear. Consider four quite different entities that might be

[1] Scriven (1978) has made a similar distinction except that he uses the terms merit and value in the same sense that we use merit and worth. We believe the latter formulation to be superior because it avoids confusion between the terms value and evaluate. Value seems to be a more generic term in this context; the placing of value is waht results from the evaluation process. We agree with Scriven, obviously, that there are two ways to value; we prefer to avoid the redundancy and confusion engenerdered when one of the subtypes is called by the same name as the more generic type.

evaluated: a professor, gold, a mathematical proof, and a language arts curriculum. A professor might be judged for merit on his or her scholarliness, i.e., his or her standing in the academic community of peers and his or her relative contribution to the discipline. Obviously merit so determined would accompany the professor wherever he or she happened to be employed; scholarliness is an intrinsic characteristic of the person. But her or she might be judged for worth on such factors as: that, as a woman, she provides a good role model; that she rates high on entreprenuership, bringing in many thousands of dollars annually in outside grants; or that she teaches in a high-demand field, that is, one in which student enrollments are burgeoning. Worth depends on what is thought to be important in the local context; a college already having many women on its faculty, that cared little about attracting outside funds, and that maintained an upper limit on its enrollments would not find this particular professor especially worthy.[2]

Gold might be judged for merit on its inherent beauty--it is attractive, and also has properties that permit it to be easily fashioned into objects of beauty, such as jewelry, plate, or statuary. It is judged for worth, however, in the trading marts of London, Zurich, and Frankfort, and daily fluctuations in it worth are worthy subjects for newspaper accounts worldwide. As you are well aware, the current worth of cold is well in excess of $300 an ounce.

A methematical proof might be judged for merit on its parsimonious-

---

[2]A little though reveals that with respect to the evaluation of a professor, it is primarily merit which is at stake in a promotion decision and worth in a tenure decision. Most universities continue to confuse these two concepts and make tenure and promotion decisions virtually equivalent in terms both of criteria and decision processes. See Scriven (1978) for an excellent discussion of this problem.

ness and its elegance. Given two proofs of the same proposition, the shorter is preferred, and a direct proof is considered more elegant than an indirect one. Boolean algebra, when first proposed, was greatly admired not for its utility (that came later) but for its inherent logic and coherence; it was, in a word, an elegant formulation. That same proof, however, is judged for worth on its applicability in a practical context. The mathematics of complex (imaginary) numbers, for example, was considered a mere mental toy until it proved useful in solving the equations of alternating electrical circuits which had, until that time, withstood conventional analysis.

A language arts curriculum might be judged for merit on its simplicity, e.g., its straightforward form, targeted scope, lack of convoluted writing, and its degree of integration, that is, the extent to which all of its substantive and formal components are properly articulated or "fit." But it is likely to be judged for worth on such criteria as the extent to which it produces student learning, it appropriate to the ability level of the students with whom it will be used, is free from bias (especially bias toward the sex, ethnicity, national origins, religion, or culture of the students exposed to it), and is teachable by the average teacher in an adopting system.

Merit and Worth as Pluralistic Phenomena

Since both merit and worth are statements about different aspects of values, it is of interest to inquire whether their estimates, once determined, are fixed, or whether they can be altered as the result of any particular circumstances. Does the value of an object, in either the sense of merit or worth, remain constant, or can either merit or worth, or both, shift from time to time or place to place?

In one sense, both merit and worth are variable. While merit is an estimate of intrinsic value, and would therefore seem to be an immutable property of the entity whose merit is being assessed, it is clear that persons competent to judge merit, for example, academic peers judging scholarliness or mathematicians judging the elegance of a proof, may differ among themselves both in stating indicators of merit and in assessing the merit of any particular entity in terms of those indicators. Just how does one tell whether scholarliness characterizes any professor, and how does one assess Professor X on the possession of those characteristics? Different responses to these questions will lead to substantially different judgments of merit.

Similarly, one may ask how one determines indicators of worth and assesses a particular entity in their terms. What does it mean to be a good role model, in the case of a professor, and how can one tell whether a curriculum is appropriate to the learning ability of the youngsters exposed to it? How does one tell whether this particular language arts curriculum possesses whatever characteristics one believes are found in good curricular? Again, judgments about both indicators and the degree to which Curriculum X conforms to those indicators are likely to vary depending upon who makes the judgments.

From this analysis one might conclude that merit and worth are equally variable. But there is a sense in which judgments of worth are a great deal more variable than are judgments of merit. That contingency arises from the fact that judgments of merit are tied to intrinsic characteristics of the entity itself, but judgments of worth depend upon the interaction of the entity with some context. A professor, whatever his or her merit, has different worth depending on whether that worth is being judged in the context of a university

seeking a scholar, one seeking an entrepreneur, or one seeking someone who can teach several sections of beginning business law. The worth of gold, whatever its beauty, depends on what speculators are willing to pay for it on any given day. the worth of a mathematical proof, whatever its elegance, depends on the existence of a real-world problem (which might be another mathematical problem) which it can help solve; Boolean algebra was relatively worthless, although extremely meritorious, until its relationship to verbal logic was exploited. And a language arts curriculum, however simple and wellintegrated, is relatively worthless with a group of black ghetto youngsters if it was designed for use with an upper middle class white clientele.

What is important to note in all these instances is that while merit remains more or less constant, at least in the sense that it is not unreasonable to expect that consensus about any object's merit can be reached, worth can and does change dramatically: change the context and you change the worth. There are several immediate and crucial consiquences of this fact:

1. The determination of worth requires an ad hoc evaluation at every site in which the entiry being evaluated is contemplated for use. It is impossible to warrant an entity for use generally, once and for all. To be sure, it is possible to conduct an evaluation to determine the worth of the entity in a number of different situations, and then to describe the entity's worth in each of those different contexts. But in that case, before worth in Context X could be inferred, one would need to have a great deal of descriptive information (often called "thick description") about the settings in which the evaluations had taken place, and about Context X, so that the degree of fit between X and

any evaluated context could be determined. Thus, local, specific evaluations to determine worth cannot be avoided.[3]

2. One very important characteristic of local settings is the values held by the several audiences--stakeholders--to whom an evaluation may be addressed. It has become patently clear over the past decade and a half that the United States is not a value consensual society but a pluralistic one. Hamilton (1977) has stressed the fact that so-called responsive models of evaluation have appeared in recent years in partial response to the recognition of this country's pluralistic value structure. Worth must be assessed, at least in part, in relation to those differing values. It must of course also be assessed in relation to other contextual factors, e.g., the nature of the student body, the qualifications of the teachers, the extent of available resources, and the like.

3. The fact that pleualistic values exist leads to the inescapable conclusion that a particular entity may be judged worthy by one group of stakeholds but worthless by another group holding conflicting values. In general, there can be no consensual determination of the worth of an entity. And just as there can be different assessments within a particular setting because of different values held by important local reference groups, even more so is it likely that different assessments

---

[3]This fact has enormous implications for innovation dissemination strategies based on the assumption that innovations can be engineered, tested, warrented for general use, and disseminated to potential adopters. Each adoption situation must be assessed to determine whether the innovation fits the situation; the mere fact that it has fit in some situations is no guarantee that it will fit in this situation.

will be reached between sites because of different values held in each. Thus, while there may be differences among stakeholders in, say, Peoria, Illinois, those difference are likely to be smaller than the differences between Peoria and Birmingham, Alabama, because these two exist in rather different cultural settings. It should be noted that these different judgments of worth, even conflicting judgements of worth, may all exist despite the level of merit of the entity; meritoriousness does not lead automatically to judgments of worthiness.

4. Methodologically, the most important consequence of the preceding deductions is that evaluations of worth must be grounded in field studies of local contexts.[4] While an experienced evaluator may be able to generate a priori hypotheses about the value (and other contextual factor) positions in a given setting, in most cases the evaluator will want, at the very least, to check his or her hypotheses, if not generate them de novo, by close study of and interaction with the context itself, especially with the people who inhabit it. And field studies call, in general, for naturalistic approaches supported by qualitative methods, as in anthropology; rather than for scientistic approaches supported by quantitative methods, as in experimental psychology.

The Determination of Merit and Worth

Merit, an intrinsic property of the entity being evaluated, is determined in one of two ways: by assessing the degree to which the

---

[4]The concept of grounding has emerged as crucial to inquiries generally. See Glaser and Strauss (1967) for a thorough explication of this term.

entity conforms to certain standards upon which a relevant professional group or group of experts agree, which might be called absolute merit evaluation, or by comparing the entity to other entities within the same class, which might be called comparative merit evaluation. So for example, in determining the merit of a professor, we might canvass a group of professors about what they take to be meritorious characteristics; whatever emerged, e.g., scholarliness, would then be used as a basis for testing the performance of actual professors. Or, in the case of a language arts curriculum, we might canvass a group of language arts educators, writers, editors, and so on, to determine what the characteristics of a meritorious curriculum are. A proposed curriculum could then be tested against the named criteria. Both these instances are examples of absolute merit evaluation. But we might in many practical situations be more interested in comparative judgments. Given that there are three professors to be considered for merit increments this year, but that there is enough money only for one increment, we may not care what each professor's absolute standing is on scholarliness but might rather wish to focus on their relative scholarliness in order to be able to decide among them. Or, in the case of curricula, we may not care what the absolute degree of integration of a curriculum is but only whether it is more or less integrated than another curriculum currently in use.

Worth, an extrinsic property of the entity being evaluated, is determined by comparing the entity's impact or outcomes relative to some set of external requirements, e.g., the results of a needs assessment or a context evaluation. What is of interest is an assessment of the entity's benefits with respect to a set of criteria. These criteria

are not drawn from a professional group or a group of experts but from the variety of local stakeholding groups that are related to or affected by the entity being evaluated. So, while merit criteria may be relatively stable, criteria of worth are highly variable depending upon which stakeholding group is being assessed, and in which context. Thus, the minority faculty in an institution may insist that newly recruited faculty must provide good minority role models, while the Vice-President for Research and the Vice-President for Administration may be much more concerned about their ability to attract outside grants. The NAACP in a local community may be more concerned about whether a proposed curriculum is free from cultural bias than are the teachers, who may instead focus upon the curriculum's appropriateness for the learning levels of their students and on their own competence to teach using the new materials and techniques.

There is of course no guarantee that any statement of criteria of worth that might be made by a stakeholding group represents its actual operating position. Scriven (1978) in a discussion of merit vs. value (worth),[5] as these relate to faculty evaluation in a university, notes that four different kinds of value (worth) criteria emerge:

- o Alleged or rhetorical values: the value system to which an institution publicly subscribes. So for example, "Teaching is very important at Berkeley--we never take personnel action without evidence about teaching." (p. 23)

- o actual or true values: the values that may be deduced from the institution's actual practices. "Solid evidence of first-rate research if required, only weak evidence of minimal teaching performance is required." (p. 23)

[5] See Footnote 1, above.

- o interests of the institution: the "set of factors the promotion tion of which would actually be valuable for--beneficial to the interests or welfare of--the institution (a really massive commitment to improving undergraduate education to the point of hiring (promoting, tenuring) for teaching talent, with research ignored, the use of a serious process of evaluation for teaching, etc.)." (p. 23)

- o ideal values: the normative value system that the institution should have. "There is the possibility that the morally, legally, socially and/or educationally correct set of (personnel) values for Berkeley is the research set." (p. 23)

In general, we would argue, the evaluator should avoid being taken in by alleged or rhetorical values, study practices to determine the actual or true values, determine insofar as he or she can what is in the interest of the institution or stakeholding group that holds those values and convey that information to them, and finally, determine whenever possible a normative set of values for the situation and compare the actual or true values to it. The latter step may not be possible in many situations because reference to "ideal" values must at once raise the question, "Ideal for whom, or in what sense?" Answering this question may be well beyond the scope of the typical evaluations.

The Relationship of Merit and Worth

In practice at least three different ways may be found in which merit and worth are linked:

Case 1: Merit and worth taken to be synonymous, that is, an institution or stakeholder group elects to define the two a equivalent, so that no benefits are taken into account. Scriven (1978) apparently beleives this to be the case at Berkeley for personnel decisions; thus both promotion and tenure at that institution seem to be dependent

upon assesements of research and scholarly productivity. Many institutions make this definition by default when, in their faculty handbooks, they treat promotion and tenure decisions as equivalent and pose precisely the same criteria (usually teaching, research, and service) for, both. Merit and worth are often defined as equivalent by default in dissemination programs for Federally funded innovations, for it seems to be assumed that a meritorious innovation can be warranted for general use in the schools (recall for example the "proven products list" of the Office of Education or the more recently established Joint Dissemination Review Panel's effort to certify innovations that have been evaluated positively).[6] In either case, it is apparent that the definition of worth as equivalent to merit is probably an error; the distinction is worth preserving.

Case 2: Merit and worth declared to be completely independent.

In a tenure case, for example, it might be argued that benefits derived from the presence of a certain faculty member (as a minority group role model, say) are so powerful that it really does not matter what his or her merit is. In the most difficult days of affirmative action there seems to be little doubt that many universities took just this posture. In the case of a currcular innovation, a local school system, under great public pressure for accountability and anxious to do something to demonstrate its commitment to change, might well jump on the bandwagon regardless of the merit of the innovation. Obviously the long-run negative effects of such posutres are serious; no institution or reference group can long take such a posture and survive.

---

6 It should be noted that the National Diffusion Network, also Federally supported, takes the posture that local school systems should be encouraged to look at JDRP approved innovations that appear to be responsive to some locally identified need and then to make whatever adaptations are necessary.

It is also possible to define merit as independent of worth. University faculty members may stretch to define a position for which staff recruitment is to occur in such a way that institutional benefits are overlooked for the sake of getting a "big-name scholar." Similarly the worth of an innovation in a local setting may be downgraded, virtually to zero, simple because the proposed innovative curriculum has "gotten such good grades in national evaluations."

As Scriven (1978) suggests, merit standards drive out standards of worth in some situations, while in other cases standards of worth drive out merit standards. Neither situation is desirable.

Case 3: Worth defined as consisting of some minimal level of merit plus other benefits. Without the minimum level of merit, worth is automatically zero. Thus, a professor who was not minimally meritorious (as demonstrated, say, by the fact that he or she was continuously promoted) would not be considered for tenure regardless of what other benefits might accrue to the institution by virtue of his or her presence. But minimal merit would not be enough; well-defined benefits would also have to be demonstrated to lead to a positive tenure decision. In the cast of a curriculum, local adoption would not occur for a porgram that was not minimally meritorious, but adoption would also not occur unless other benefits could be demonstrated as well.

It is also the case that decisions about merit are usually not made in the absence of simultaneous considerations of worth. Universities do not usually hire a professor simply on merit; some possibility that this professor could also serve some university purpose muust also exist. And no one would undertake to develop a curriculum no matter what its

promise on merit unless he or she had reason to believe that the curriculum would also be useful in some setting.

The methodological problem posed for the evaluator in this more typical real world case (Case 3) is to avoid becoming confused about the separate components that go into "balanced" judgments simply because they are interrelated. While merit decisions reinforce decisions of worth, and vice-versa, they are separate decisions, made on separate criteria, and require different methodological approaches to be established.

The Relationship ofMerit and Worth To Formulate and Summative Evaluation

The relationship of merit and worth to the concepts of formative and summative evaluation (Scriven, 1967) is complex. It is easy to be misled into believing that merit evaluations are formative and worth evaluations are summative. The aim of formative evaluation is, after all, refinement and improvement, which makes one think of intrinsic aspects of entities that can be refined and improved. The aim of summative evaluation is to determine impact or outcomes, which makes one think of contexts in which such impacts or outcomes may be noted. But in fact the dimensions of merit/worth and formative/summative are orthogonal; evaluations of merit can be either formative or summative just as can evaluations of worth.

This assertion can perhaps be best understood through an example. We turn by way of illustration to the case of an innovative program or curriculum which is to be developed by a national team and then to be considered for adoption in a local setting--a quite common

situation. Evaluations that seek to establish the merit of such an innovative program or curriculum might properly be termed developmental evaluations--evaluations occurring during the process of development that seek to improve or refine the entity to its optimal state as an instance of its kind. Evaluations that seek to establish the worth of such an entity may be thought of as adoptive evaluations--evaluations occurring after a completed entity summatively evaluated for merit is available that relate to the use of the entity in some concrete situation. Both developmental and adoptive evaluations may be carried out for either summative or formative purposes; four types are determined by crossing the merit/worth and formative/ summative dimensions. We identify and analyze these four types below in terms of their purposes, their audiences, and the sources from which judgmental standards are drawn:

1. The purpose of formative developmental evaluations is to modify and improve the design of an innovative program or curriculum while it is still under development. The audience for such an evaluation is the national developmental team, since they are the agents who can act on the evaluation information to make whatever changes are indicated. The sources of standards are panels of professional peers or other expert groups, who have reason to be well acquainted with the characteristics of meritorious innovations of that type, e.g., curriculum experts, substantive experts, substantive experts, other evaluators, and the like. While such groups no doubt will not agree on every detail they are most competent to specify what shall be taken into account in determining merit.

2. The purpose of summative developmental evaluations is to critique a completed entity in terms of professional or expert standards so as to be able to certify and warrant merit. The primary audience for this evaluation is the group ofpotential adopters who need to be reassured about merit before they can reasonably consider the worth of the innovative program or curriculum in their own contexts. The source of standards continues to be panels of professional experts.

3. The purpose of formative adoptive evaluations is fit or adapt the program or curriculum to a local context or situation. As we have noted, there will always be local contextual, including value, differences, and no innovation could hope to fit all contexts. What is required is a local assessment of context and values so that the fit of the innovation can be determined. Refinements, adjustments, and other adaptations will be called for to optimize the fit; a local adaptation team is the proper audience for formative adoptive evaluation informations, and the local assessment of context and values (which might be construed as a needs assessment) becomes the source for standards.

4. The purpose of summative adoptive evaluation is to certify or warrant the adapted program or curriculum for permanent local use. The audience for this evaluation is that group of local decision makers (probably some from each of the stakeholding audiences) which will make or shape the final decision about whether to permanently adopt the innovation. This group will vary widely in composition and clarity from place to place depending upon how decisions happen to be made in each locale. The source of standards for this evaluation is a local needs assessment, for unless the innovative program or curriculum meets local needs there can be no justification for adopting it.

This example focuses on the adoption of an innovative program or

curriculum, but the distinctions which are illustrated by it apply whenever merit and worth are at issue. In the case of the University professor, as a further example, one can imagine a formative merit evaluation in which the professor is himself or herself the audience (the professor is in the final analysis responsible for his or her own behavior), the purpose is to provide feedback that will help him or her to improve in the sense of becoming more meritorious, and the standars are the professional standards of his or her own peer group, other educational psychologists, say. Many universities now provide for annual reviews, sometimes in connection with awarding salary increments, that serve this function. The summative merit review occurs when the professor is formally considered for promotion in rank. Summative worth reviews are made when the professor is considered for tenure, and formative worth reviews can also be regularly scheduled (although they usually are not) to provde the professor feedback about how he or she is doing with respect to Unversity criteria or worth (e.g., role modeling, entrepreneurship, carrying a fair teaching load, and the like). Similar examples could be provided in virtually every evaluation situation.

Summary

In this paper we have attempted to distinguish between two kinds of value determinations: Those of merit and of worth. Merit, we have suggested, is an intrinsic measure of value which inheres in the evaluated entity with its context and thus may vary dramatically from context to context. We gave examples of both merit and worth in relation to four evaluated entities: a university professor, gold, a mathematical proof, and a language arts curriculum.

We pointed out that while both merit and worth are somewhat variable worth is especially variable because of its unique dependence upon the context in which it is assessed. As a result of this fact, we suggested, four important consequences resulted: the determination of worth requires a separate, ad hoc evaluation in each proposed context, that is, that worth cannot be determined once-and-for-all; that the values of stakeholders form an important part of each local context; that an entity may be judged worthy by one group or subgroup but completely unworthy by another, because of value conflicts; and that (and from our point of view, most important) evaluations of worth must be grounded in field studies of local contexts and cannot be carried out by the conventional experimental, pre-post inquiry modes that have dominated evaluation in the past.

We went on to suggest that merit may be determined by absolutely, in respect to judgments made by professional peers and experts, and comparatively, in respect to other entities of the same type as that being evaluated. Worth is determined in relation to a set of benefits that might be derived, as stipulated by stakeholding audiences. We suggested a caveat, however: that stated positions of stakeholders were not necessarily their true positions, and that the evaluator needed to keep these two sets (as well as institutional interests and ideals separate.

Merit and worth were found to be related in practice in one of three ways. They are sometimes taken as synonymous terms, and, at the other extreme, are often taken as complately independent. A more realistic and meaningful posture consists in defining worth to depend on some minimal level of merit but to include, in addition, certain other benefits defined differently in each local context.

Finally, we related the merit/worth dimension to the classic formative summative dimension, found them to be orthogonal, and so generated four types of evaluation. These were illustrated in the cast of an innovative program or curriculum and found to differ significantly along dimensions of purpose, audience, and sources of standards. It was asserted that similar analyses could be made of the four types in virtually every evaluation situation.

The analyses and arguments we have made here have been brief but, we feel, compelling. The distinction between merit and worth is not only conceptually intriguing but operationally heuristic. If taken seriously, the differences noted have enormous implications for both the theory and the methodology of evaluation, particularly, we believe, in pointing toward the need for more naturalistic and less experimentally oriented evaluation.

Note

This paper is based on Chapter 4 of a book in preparation by the present authors tentatively entitled Naturalistic Evaluation to be published by Jossey-Bass Publishers in 1980.

References

1. Glaser, Barney G.; and Strauss, Anselm L. The Discovery of Grounded Theory. Chicago, IL: Aldine Publishing Company, 1967.

2. Hamilton, David. Making sense of curriculum evaluation: continuities and discontinuities in an educational idea. In Shulman, Lee S. (Ed.) Review of Research in Education. Itasca, IL: Peacock Publishers, Inc., 1977, 318-48.

3. Scriven, Michael. The methodology of evaluation. AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago, IL: Rand McNally, 1967, 39-83. Also reprinted in Worthen Blaine R., and Sanders, James R. (Eds.) Education Evaluation: Theory and Practice. Worthington, OH: Charles A. Jones Publishing Co., 1973, 60-104.

4. Scriven, Michael. Merit vs. value. Evaluation News, No. 8 (December, 1978), 20-29.