

https://dspace.org/

User-focused evaluation of training programme effectiveness in a South African industrial company.

Cole, Mary Berrington

2015-02-13
http://hdl.handle.net/10539/16980

Downloaded from a DSpace repository. https://dspace.org/

He points out that to date most evaluation methodology has focused on inquiry and has come from a narrow range of disciplines in the social sciences.

Citing Guba, Smith (1981a, p52) goes on to float the idea of metaphors for evaluation, as one means of developing new methods, as well as a means of communicating about methods with evaluation practitioners and theorists.

> A metaphor, as commonly defined, is a figure of speech in which the meaning of a term or phrase is transferred from the object it ordinarily designates to another object so as to provide new insight or perspective on the latter. In a broader sense, the term metaphor, may also designate a process whereby the meanings and relationships of one theory or model may be used to suggest meanings or relationships in another arena.

The usefulness of metaphors and metaphor-like conceptual translations is well established as a heuristic (thinking pattern) device, as the writer can bear out from her own frequent use of this technique. The classical or dominant evaluation paradigm is already graphically identified with the agricultural-botany research tradition, where subjects are likened to seedlings, in treatment groups or experimental plots, with treatments being likened to the application of controlled amounts of water, fertiliser, etc., and the treatment outcomes likened to plant size, yield and so on. Similarly, the alternative evaluation paradigm is closely identified with the field anthropologist investigating tribal culture. Evaluation models presently identified with other metaphors are:

illuminative	- shedding of light on...
systems analysis	- micro-economics
countenance	- different facial visages
adversary	- counsel for the prosecution and defence
etc. (see section 2.3)	

Smith (1981a and 1981b) is advocating that the time has now come to look further afield for metaphors for evaluation and provides examples of investigative reporting, law, architecture, geography, philosophy, literacy/film criticism, and

watercolor painting, as metaphors for evaluation. He provides a category system for reviewing potential metaphors (Smith, 1981a, p56). Clearly, there is ample scope for evaluators to creatively expand their repertoire of metaphors and new developments in this field hold promise of the emergence of exciting new tactics for creative, problem-solving evaluation.

Creative evaluation means working within the paradigm of choices so that different situations can be approached without preconceived ideas about which particular evaluation methods or approaches ought to be used. Creative evaluation is therefore a problem solving approach which emphasizes adaptation to different conditions. This contrasts with a technical approach which attempts to mould and define conditions to fit preconceived ideas about how an evaluation should be designed and implemented.

Evaluators are trained to be critical thinkers as opposed to creative thinkers. Creativity in evaluation is critical if the promise of the new discipline of the new evaluation is to be fulfilled. The practice of normal science in the social, behavioural or anthropological sciences is quite different from the practice of creative evaluation. The real problem with the reliance of evaluators on paradigms and heuristics (patterns of thinking) is our general lack of awareness of the extent to which they create barriers to creative thinking. It is difficult to be attuned and responsive to the uniqueness of each new evaluation situation when a scientific paradigm is controlling the analytical process, screening unfamiliar data and locating the new experience within the narrow parameters of our past experience, making available primarily those approaches used most often in other situations. Evaluators are typically not aware that what may appear to be discretionary judgements are really predictable routines. Patton maintains that it is possible for evaluators to become alerted to their paradigms and heuristics and thereby release their creative potential.

Creative evaluators are keen observers of their own actions and ways of thinking, as part of the ongoing process of observing and studying the evaluation exercise. They use their observations to help keep themselves situationally responsive, open to new possibilities and capable of reaching into their large repertoire of techniques, methods and approaches to find or adapt a technique, method or approach that is appropriate to the situation at hand.

Patton (1982, p300) also includes in his approach the requirement for evaluators to be possessed of fundamental skills

> ...that are different from those of any single social or behavioural science and broader than the basic skills of academic scholarship. Just as the new standards of evaluation are different from the traditional standards of the academic disciplines, so too the professional skills needed by practicing evaluators (at least those who take the new standards seriously) are different from the scholarly skills of basic sciences.

In view of this, he believes that evaluators ought to be capable of playing multiple roles, including that of evaluation trainer. Patton regards every evaluation as a training exercise, to help participants in the evaluation (decision makers, stakeholders, information users, programme staff, funders and so on) understand what evaluation is all about. An evaluator can teach by example or by following a specific course of training.

Patton argues that the new standards require evaluators to play multiple roles and take on multiple responsibilities. Previously evaluators were only held responsible for the production of accurate information and new knowledge. How it was used was not their concern.

> Today, the evaluator is called on to take responsibility for how evaluative information is used, to take responsibility for the correct implementation of evaluation procedures, to be responsible for the protection of human subjects in evaluation research, as well as the traditional responsibility to produce accurate information.

> The eminent humanistic psychologist Carl Rogers uses a metaphor with his students that dramatically captures this change in responsibilities. In the past scientists have been like ammunition wagons, loaded with powerful knowledge just waiting to be used, and available to either the good guys or the bad guys. Now, applied social scientists, particularly evaluators, are called on to target their knowledge and take

responsibility for what they "hit." Rogers (1961, cited in Patton, 1982, p 305) puts it quite forcefully: "Don't be a damn ammunition wagon. Be a rifle."

2.2.5.1.11 Definition and Procedure

Reviewing Patton's framework for evaluation in the light of the evaluation literature, it becomes apparent that evaluation which addresses the information needs of specific people associated with a training programme, does indeed have a significant place in the evaluation literature. This user-focused, or utilization-focused evaluation, is defined as follows:

> The practice of evaluation involves the systematic collection of information about the activities, characteristics, and outcomes of programs, personnel, and products for use by specific people to reduce uncertainties, improve effectiveness, and make decisions with regard to what these programs, personnel or products are doing and affecting. This definition of evaluation emphasizes
> (1) the systematic collection of information about
> (2) a broad range of topics
> (3) for use by specific people
> (4) for a variety of purposes (Patton, 1982, p35).

The steps in user-focused evaluation are:

> identify and organize relevant decision makers and information users;
>
> the relevant evaluation questions are identified and focused;
>
> evaluation methods are selected that generate useful information for identified and organized decision makers and information users;
>
> decision makers and information users participate with evaluators in data analysis and data interpretation;
>
> evaluators and decision makers negotiate and co-operate in dissemination efforts (Patton, 1978, pp284-288).

Patton (cited in Smith, 1982a, p396) added that

> ...perhaps the most surprising reaction to (Utilization Focused Evaluation), and the reaction that is most often expressed in mail I get about the book, concerns its legitimating function for evaluators who had developed a "utilization-focused" style of work, but hadn't quite articulated, systematized, or labeled what they were doing in a way that allowed them to be comfortable that what they did was legitimate.

This is precisely what Patton's work has done for this author - it legitimated an approach to evaluation she had developed intuitively, experientially and by trial-and-error.

* The second edition of Utilization Focused Evaluation (Patton, 1986) was received too late for inclusion in this analysis of Patton's view of evaluation. The new edition treats the qualitative-quantitative debate quite differently (see pages 209-217).

2.2.5.2 Discussion

The review of the new evaluation literature is discussed in terms of the nature of evaluation utilisation. Simultaneous with Patton's work, Alkin, Daillak and White (1979) have addressed the importance of evaluation utilisation. They query why we should even be concerned with this question and find the answer in the fundamental distinction between 'evaluation' and 'research' which is essential to the development of evaluation theory and ultimately to the practice of evaluation. "On the one hand, there are studies designed primarily to add to the body of knowledge (research), on the other, those studies designed primarily to provide information for decision-making (evaluation). And these two functions are separate and distinct" (Alkin, 1973). "The basic rationale for evaluation is that it provides information for action" (Weiss, 1966). Viewed from this perspective, evaluation is distinctly practical and utilitarian, and to succeed, or work, it must have influence upon the decision-making process in its target programmes.

Alkin et al (1979) identify two divergent views on the influence of evaluation. One, the mainstream viewpoint, contends that evaluations seldom influence programme decision-making and is pessimistic about evaluation ever achieving a break-through to really have an impact on programmes. Alternatively, there is a minority point of view in the literature which concludes that evaluations do already influence programmes in important and useful ways.

Alkin et al (1979) point out that two of the very few existing empirical studies of evaluation utilization conducted in field settings tend to support the view that evaluation can, and often does, influence program decision-making. Alkin himself, and associates at the Center for the Study of Evaluation, University of California, Los Angeles, examined the impact of evaluation on decision-making in 42 federally funded special education programmes. The second study was conducted by Michael Patton and associates at the Minnesota Center for Social Research. These researchers looked at 20 evaluation studies of national health programmes, and for each programme studied interviewed the evaluator and a major program decision-maker. In the Alkin et al (1979) study, which looked at evaluations being conducted at the local level, and in the Patton, Grimes, Guthrie, Brennan, French and Blyth (cited in Weiss, Ed., 1977) study of evaluations that were national in scope, evaluation is found to be used successfully for program evaluation, and as having an incremental, rather than a decisive effect on programme decision-making.

What is the success factor in the evaluations documented by Alkin and Patton? Significantly, the researchers were able to identify a new, unanticipated factor, which they termed the personal factor.

> The personal factor emerged most dramatically in our interviews when, having asked respondents to comment on the importance of each of our eleven utilization factors, we asked them to identify the single factor that was most important in explaining the impact or lack of impact of that particular study. Time after time the factor they identified was not on our list. Rather, they responded in terms of the importance of individual people....
>
> The specifics vary from study to study but the pattern is markedly clear: Where the personal human factor emerges, where some individual takes direct, personal responsibility for getting the information to the right people, evaluations have an impact. Where the personal factor is absent, there is a marked absence of impact. Utilization is not simply determined by some configuration of abstract factors; it is determined in large part by real, live, caring human beings (Patton et al cited in Alkin et al, 1979, p24, emphasis in the original).

This statement reflects the key man concept.

Alkin et al (1979) believe that the conflict between mainstream and alternative perspectives on evaluation impact is, first and foremost, a conflict between definitions of utilization. Most commentators adopt as their criterion for utilization immediate and direct impact on one or more critical programme decisions. Judged by this cut-and-dried standard, utilization seldom occurs, since decisions are seldom made which conform to the adjudged worth of a programme. However, the Alkin and Patton studies concentrate less on this notion of dramatic impact on program decision-making and expand the definition of utilization to include gradual influence on programmes.

In her paper entitled 'Knowledge Creep and Decision Accretion', Weiss (1980, p381) says she uses the mangled terminology to make two points. The first is that knowledge derived from research findings is not often

> ...utilized in direct and instrumental fashion in the formulation of policy... Instead, research knowledge usually affects the development and modification of policy in diffuse ways. It provides a background of empirical generalizations and ideas that creep into policy deliberations. Its influence is exercised in more subtle ways than the word 'utilization' - with its overtone of tools and implements - can capture.

The second point that Weiss (1980, pp381-382) makes is that

> ...many policy actions, even those of fateful order are not 'decided' in brisk and clear-cut style. The term 'decision' implies a particular set of events. ...But in large organizations, policies often come into being without such systematic consideration. ...Yet the onrushing flow of events shape an accommodation - and a pattern of behaviour - that has widespread ramifications. It may in time be ratified by conscious policy action, but in the crucial formative stages, it just seems to happen. Without conscious deliberation, the policy accretes.

Weiss (1980, p385) believes that the fact that decisions take shape in this informal way, helps to explain the lack of direct use made of research. Staff in an organization base their ongoing actions on the sum of their knowledge and judgement.

> Even scientists, who work in a tradition that requires the citation of sources, find it difficult to trace the genealogy of their ideas. Decision makers, who are less attuned to footnoting their references, are probably less aware of the origins of many of the concepts with which they make sense of the world. Yet they suspect that some of their ideas come from the research to which they have been exposed.

Reports of general uses of research included:

- the flow of research aids policy formulation and direction setting;
- research establishes needs, both type and distribution;
- evaluation research shows what works;

- social science research is a medium of communication, especially about emerging issues - a form of continuing education;

- research use can be ritualistic providing a scientific patina to organisation documents;

- research is used to support a point of view and persuade other participants in the decision-making process, i.e. ammunition for political wars, which suggests that research has a special standing in organizational debate and bargaining;

- research is used to clarify and inform the taking of a position;

- conceptual use is made of social science research to gain insights into social processes as a foundation for understanding (Weiss, 1980, pp 386-390).

Weiss continues:

> Many people underscored the point that social science research was not their only source of information and ideas. It was one source among many, and not usually powerful enough to drive the decision process. They read social science research, but it was auxiliary to their own first-hand experience. They were exposed as well to information from colleagues, program records, meetings and conferences, books and television, and the conversation of friends. ...Those social scientists who expect research to be authoritative enough to determine policy choices are giving insufficient weight to the many and varied sources from which people derive their understandings and policy preferences... ...people in official positions often do not catalog research separately in their minds. They interpret it as they read it in light of their other knowledge, and they merge it with all the information and generalizations in their stock. ...The lessons from research have become part of a gestalt (Weiss, 1980, pp 390-391).

Weiss (1980) concludes that if reports of research use seem fuzzy, this fusion of research with knowledge from other sources is part of the reason. People do not always know the source of the ideas and models which form their basic intellectual capital or which of them came, directly or indirectly, from social science research. Finally, she points out that when people are asked how they use social science research, they are in effect being asked to link social science <u>input</u> to decision <u>output</u> and many people in organisations do not seem to recognise, or are not willing to acknowledge, that they make decisions.

> They recommend, advise, confer, draw up budgets, testify, develop plans, write guidelines, report, supervise, propose legislation, assist, meet, argue, train, consult - but <u>decide</u>? It is not a concept that seems to describe aptly the flux of activities that engage officials even on the top rungs of complex organizations (Weiss, 1980, p 392).

Instrumental utilization of social science research is not often found in organisations and has led to some conclusions that research is ignored. However, when the many and varied ways in which research can be used is recognised, then the influence of research appears much more positive. Weiss (1980) found that the people she surveyed had a strong sense that social research mattered, even if they could not site any conscious use of it. They believed that they had assimilated generalizations, concepts and perspectives that shaped their understanding and determined their actions. Social science research seems to have a pervasive effect. By knowledge creep, it provides an underlying set of ideas, models of the interaction of people, conditions and events, which enter into our images of how the organization works.

However, despite this general usefulness of social science research, there is a

> ...series of disjunctures between research and decision that limit the use of research. Research does not fit the exact circumstances within which decisions are made, research is not ready on time for decisions, research conclusions are not unambiguous or authoritative enough to provide direct guidance, research reports do not reach the right

> audience, decision makers do not understand or trust research findings or understand how to interpret and apply them, the lessons from research are outweighed by the combination of competing interests, agency, self-protection, and individual career concerns (Weiss, 1980, p 398).

On top of this people do not believe that they make decisions, because fragmentation of authority in most organizations is both vertical and horizontal. Many decisions happen without acknowledged responsibility, limitations of time and events, purposefulness, calculation or sense of significance and thus they just seem to happen or accrete.

In investigating social science research utilization, Weiss (1980) has given a yardstick whereby evaluation utilization or non-utilization should be judged. As has been pointed out by Patton, Alkin and Weiss, the basic rationale for evaluation is that it provides information for action and to be judged successful, it must have an influence on the decision-making process, by being of direct and specific value to identified staff in programmes with needs for specific information.

By conducting user-focused evaluation, the evaluator in fact makes a fundamental choice between practice based evaluation and discipline based evaluation. By addressing an audience of training programme staff in the company, the evaluation emphasises relevance, rather than the search for universal truth, and thereby becomes practice-based evaluation. This is an option which has only been open to evaluators since approximately 1975.

Previously, evaluators had no option but to follow the research methods of their own disciplines in the social and behavioural sciences. Their research methods were concerned with the discovery of knowledge which was generalisable, by means of methods which were valid, reliable and replicable. The audiences they addressed were their peers in their disciplines.

Their main concern in conducting evaluations revolved around the choice of methods. The social scientists and psychologists usually favoured experimental or objectives-based methods which yielded quantifiable data susceptible to statistical

analysis. This paradigm has been dominant in evaluation thinking and is the mould in which training evaluation has been firmly set until now. However, news from the 1984 annual conference of the American Society for Training and Development is that this traditional kind of training evaluation, usually linked to a systems model of evaluation (Hamblyn 1974, Kirkpatrick 1975, Rossi, Freeman & Wright, 1979, Tracy 1974, Warr, Bird and Rackham 1970) is now considered unworkable (Moubray J and Pansegrouw G, Directors, P-E Corporate Services (Pty) Ltd., verbal comment 1984).

The disadvantages of the pre-ordinate research designs have caused many evaluators in the education and social services programmes to choose the alternative methods of the anthropologists , with their naturalistic enquiry based on qualitative, descriptive methods. These are usually linked to responsive models of evaluation such as case study (Guba & Lincoln 1982) and illuminative evaluation (Parlett and Hamilton, 1977). There are now fairly numerous instances of the application of alternative evaluation methods to training evaluation (for example: Galvin, August 1983; Salinger and Deming, August 1982, Connally, October 1983; Clement and Aranda, August 1982). However, in terms of a new approach to training evaluation they do not serve to:

> ...construct a general framework that goes beyond the grab-bag approach of technique mongering that leaves one with a limited bag of tricks, but no real foundation on which to build new approaches to deal with new situations and unanticipated problems (Patton, 1982, p18).

Furthermore, although more flexible and adaptable, these kinds of evaluation nevertheless still have the weakness of being addressed primarily to audiences outside the programmes being evaluated.

During the last decade, problems arose concerning the usefulness of much research. Evaluation was seen to suffer from under-utilization. Evaluators began to examine their own practice with a view to establishing standards for good evaluation. These standards were published by the Joint Committee on Standards for Evaluation in 1981. They require that evaluations should be utility-focused (useful), feasibility-conscious (practical), propriety-oriented (ethical) and finally, accurate, in <u>that</u> order. Finally, user-focused evaluation requires new, creative,

problem-solving tactics for evaluation as opposed to traditional, analytical, technical evaluation. Nevertheless, it is the latter which remains the norm in training evaluation.

It is surprising that such a paradoxical situation appears to exist between training evaluation and the new evaluation. Training evaluation, as presently practised is considered impractical and therefore impossible to implement. The new evaluation offers an approach to training evaluation which would overcome these problems. Yet a transfer of thinking about evaluation does not appear to have taken place. Professor Donald L Kirkpatrick, doyen of training evaluation, has denied knowledge of the new evaluation in relation to training in industry (Kirkpatrick, D, personal communication, 1983). A citation search of Michael Q Patton since 1975 reveals that he has been cited more than 250 times, which is a record for any author in the experience of the University of the Witwatersrand On Line Service (Social Science Citation Index and Science Citation Index).* However, with the exception of Brinkerhoff writing in Training and Development Journal (Brinkerhoff, December, 1981 and August, 1983), there are no other known citations of Patton in the training and development literature. This suggests that the application of the new evaluation approach to the evaluation of industrial training could be a new and improved way of evaluating training in industry.

2.2.6 Summary

In the preceeding sections it has been shown that the evaluation literature manifests a chronological development of fundamental nature. In the period prior to 1975 sociological, phsycological and anthropological methods were applied to evaluation problems and how to evaluate was of major concern. From 1975 onwards, evaluation has begun to develop into a research field of its own and the major focus of attention has been on what to evaluate, with how to evaluate a matter of secondary concern. What to evaluate has been addressed in some forty evaluation models which serve as aids to focus ideas about evaluation. Their emergence has been associated with the development of a vocabulary, definitions, typologies and so on - all the accoutrements of a new research field. Most recently of all the

* The main database for management journals, ABI/Inform Data Base, does not have a citation index.

debates is now centering on the issue of why evaluate and who to evaluate for. This is a concern shared with research at large. In other words the focus has moved from tactics to strategy.

Training evaluation has taken its place in the evolution of evaluation but is in danger of becoming fixed in a fossilized form. This is because training evaluation is still pre-occupied with the how-to-evaluate debate, and then only within the experimental or objectives-based methods. Anthropological methods have barely been tried. As for what to evaluate, trainers have only one evaluation model in mind - systems evaluation. Finally, concerning why evaluate, there is no coherent thinking by trainers at present - only vague dissatisfaction and confusion with the existing situation.

There is a big need to transfer to training much of the new evaluation thinking which has developed on the education and social welfare side of evaluation. At present, the two sides remain separated to a remarkable degree, physically as well as theoretically. Training evaluation and education/social evaluation are categorised separately in different parts of the library or even in different libraries and training journals are quite distinct from education and sociology journals. It is high time for some bridge building between training and education at least, and this research represents such an attempt, in the field of evaluation.

2.3 Meta-Evaluation Literature

2.3.1 Introduction

Meta-evaluation is equated with evaluation of evaluation, secondary evaluation, evaluation audit, evaluation critique and evaluator self-criticism (Patton, 1982; Cook & Gruder 1978). The underlying concept is that evaluations should be subject to critical review. The authors which have made the most important contribution to the meta-evaluation literature are Scriven (1974), Stufflebeam (1974), Cook and Gruder (1978), Gowan and Millman (1978), the Joint Committee (1981) and the ERS Standards Committee (cited in Rossi, Ed., 1982). Their meta-evaluation methodology shows an evolution over time from the first proto-types for meta-evaluation of Stufflebeam (1974) and Scriven (1969) to the most up-to-date thinking of the Joint Committee (1981) and the Evaluation Research Society (cited in Rossi, Ed., 1982).

Only the latter two are reviewed here. Stufflebeam (cited in Rossi, Ed., 1982) considers that the Evaluation Research Society standards are highly consistent with the Joint Committee standards and has queried the need for independent standard-setting efforts. However, their separate existence does confirm certain widely shared principles which have been established through independent efforts. These principles provide something more than a group of disparate meta-evaluation techniques with which to view the evaluation or process issues of the research.

2.3.2 The Standards for Evaluations of the Joint Committee on Standards for Educational Evaluation

To date, the most comprehensive, carefully developed, objective, and useful way of judging evaluation plans, processes, and results, has been the evaluation standards developed by the Joint Committee (1981). The joint AERA/APA/NCME* Committee on Test Standards, which had developed the Standards for Educational and Psychological Tests, (1974) had considered including a section on programme evaluation in that document. They decided however, that such an effort was not within the scope of their task and recommended that a new committee be appointed to develop a companion volume of standards for educational evaluation generally. This recommendation was adopted by the three organisations. Subsequently, additional organisations were invited to join them and the Joint Committee on Standards for Educational Evaluation was formed. The Committee was motivated by the need for a set of professional standards which could play a vital role in upgrading the practice of educational evaluation, especially as no adequate standards for evaluation previously existed; therefore, they undertook to perform a needed service by developing and publishing such standards.

* AERA - American Educational Research Association
APA - American Psychological Associations
NCME - National Council on Measurements used in Education

A significant policy decision was the call for a committee with broad representation. There were 18 members, including 9 representing groups oriented toward improving evaluation methodology and nine representing groups oriented toward improving the application of evaluation methodology. The Joint Committee was chaired by Dr. Dan Stufflebeam, and its staff were based at the Western Michigan University Evaluation Centre. They were funded by 12 sponsors including the National Institute of Education and the National Science Foundation. By the time the standards had been drafted, reviewed, revised, field tested and presented at national hearings, several hundred educators, social scientists, education officials and other experts had been involved. They therefore cannot but be considered representative of all significant relevant opinion in the USA.

The focus of the standards is described as follows:

> The standards do not encompass a particular view of what constitutes good education, nor do they present specific criteria by which to judge educational programs, projects and materials. However, they do contain advice for dealing with these vital issues. In essence, evaluators are advised to gather information which is relevant to the questions posed by clients and other audiences and yet sufficient for assessing an object's effectiveness, costs, responses to societal needs, feasibility and worth (Joint Committee, 1981, p7).

Robert Linn (1981, p89), of the American Psychological Association, and a member of the Joint Committee, has described the standards as: "...a sensitizing device and a reminder of important aspects of an evaluation that might otherwise be overlooked. They can help an evaluator guard against proceeding with blinders to legitimate evaluation concerns while focusing on a single perspective." However, he cautions that they should not be rigidly followed as a how-to-do-it guide by the untrained or technically weak evaluator. They are not a substitute for thorough training and extensive experience.

There are 30 separate standards. They are presented in four groups that correspond to four main concerns about any evaluation - its utility, feasibility, propriety, and accuracy (See Fig. 2.4) The Joint Committee (1981, pp13-15) is satisfied that standards which shape an evaluation so that it has these four important attributes are necessary and sufficient for sound evaluation in education.

Utility category contains standards for guiding evaluations so that they will be informative, timely, and influential. The standards included in this category are Audience Identification, Evaluator Credibility, Information Scope and Selection, Valuational Interpretation, Report Clarity, Report Dissemination, Report Timeliness, and Evaluation Impact. Overall, Utility standards are concerned with whether an evaluation serves the practical information needs of a given audience.

FIGURE 2.4

Summary of the Joint Committee Standards

A. UTILITY STANDARDS
- Audience Identification
- Evaluator Credibility
- Information Scope and Selection
- Valuational Interpretation
- Report Clarity
- Report Dissemination
- Report Timeliness
- Evaluation Impact

B. FEASIBILITY STANDARDS
- Practical Procedures
- Political Viability
- Cost Effectiveness

C. PROPRIETY STANDARDS
- Formal Obligation
- Conflict of Interest
- Full and Frank Disclosure
- Public's Right to Know
- Rights of Human Subjects
- Human Interactions
- Balanced Reporting
- Fiscal Responsibility

D. ACCURACY STANDARDS
- Object Identification
- Context Analysis
- Described Purposes and Procedures
- Defensible Information Sources
- Valid Measurement
- Reliable Measurement
- Systematic Data Control
- Analysis of Quantitative Information
- Justified Conclusions
- Objective Reporting

(Adapted from Joint Committee, 1981, pp19-20, 51, 63-64, 97-98)

Feasibility, the second category, contains standards which recognize that a programme evaluation usually must be conducted in a natural, as opposed to a laboratory, setting and that it consumes valuable resources. The three standards in this category are Practical Procedures, Political Viability, and Cost Effectiveness. Overall, the Feasibility standards call for evaluations to be realistic, prudent, diplomatic, and frugal.

Propriety, the third category, contains standards which reflect the fact that evaluations affect many human beings in many ways. The eight standards dealing with Propriety are Formal Obligation, Conflict of Interest, Full and Frank Disclosure, Public's Right to Know, Rights of Human Subjects, Human Interactions, Balanced Reporting, and Fiscal Responsibility. Overall, the Propriety standards require that evaluations be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by the results.

Accuracy, the fourth category, includes those standards that determine whether an evaluation has produced sound information. The eleven standards placed in this category are Object Identification, Context Analysis, Defensible Information Sources, Described Purposes and Procedures, Valid Measurement, Reliable Measurement, Systematic Data Control, Analysis of Quantitative Information, Analysis of Qualitative Information, Justified Conclusions, and Objective Reporting. The overall rating of an evaluation against the eleven standards in this category gives a good idea of the evaluation's overall validity.

A review of the standards suggests that some are relatively unique to evaluation, while others are more general professional standards. This reflects the conviction of the Joint Committee that standards should be selected if they address persistent and important problems in evaluation, whether or not they are unique to evaluation. The standards are presented in a common form, for ease of reading and to show their interrelationships: description, standard, overview, guidelines, pitfalls, caveats and illustrative case. The standards are not intended as mechanical rules, but as guiding principles. Whatever the nature of a particular evaluation, the evaluator concerned should reflect carefully on the standards and how they apply to specific situations, especially at key checkpoints in any evaluation process.

Whether a given standard has been met in a particular situation is finally a matter of judgement. The Joint Committee (1981) does not believe that precise decision rules should be pre-specified. However, evaluators who cite these standards should communicate clearly how they were used, what judgements about meeting the standards were reached, and what evidence exists to support the judgements. They have developed a citation form on the use of the standards, for completion and inclusion at the back of the evaluation document. The citation form is shown in Appendix F.

Taken as a whole, these thirty standards are a proposed working philosophy of evaluation. They define the Joint Committee's idea of what principles should guide and govern evaluation efforts, and they offer practical suggestions for carrying them out. The Joint Committee urges that they be used by all persons who commission, conduct or employ the results of evaluations, and not just by the narrow stereotype of specialist evaluators who do evaluations for the use of others. The Joint Committee is confident that the use of the Standards will lead to sound, useful, ethical and cost-effective evaluation services.

2.3.3 The Standards for Evaluation of the Evaluation Research Society (ERS)

The ERS (cited in Rossi, ed., 1982) undertook a parallel activity to that of the Joint Committee (1981) for several reasons, including the fact that the Joint Committee's standards focus only on educational programmes. They do not therefore necessarily represent the broader interests of the Evaluation Research Society membership, which includes a wider range of programmes, including training, business and personnel systems. Nevertheless, the ERS (cited in Rossi, ed, 1982) acknowledges a considerable debt to the Joint Committee (1981) and other sources for ideas they have borrowed and incorporated in the ERS Standards for Programme Evaluation.

The ERS standards are organised into six sections: formulation and negotiation; structure and design; data collection and preparation; data analysis and interpretation; communication and disclosure; and utilisation. The standards themselves take the form of 55 simple admonitory statements within the above six sections. The standards are specific to programme evaluation and do not encompass the full body of legal requirements, governmental regulations, and accepted norms for professional and corporate conduct to which evaluators are subject (ERS cited in Rossi, ed. 1982).

Author Cole M B
Name of thesis User-focused evaluation of training programme effectiveness in a South African Industrial company 1987

PUBLISHER:
University of the Witwatersrand, Johannesburg
©2013

LEGAL NOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.