

The significance in evaluation theory and practice of the need to produce recommendations.

Author: Franz Kriven

Organisation: Office of Evaluation and Audit, Aboriginal and Torres Strait Islander Commission

Some years ago, Tony Shinkfield, who co-wrote *Systematic Evaluation* with Dan Stufflebeam, told me that Dan believed that the profession needed more practitioners to come forward and share their knowledge of how they work and get their results. At last year's AES Evaluation Conference, Keynote Speaker Elliot Stern-editor of the (International) Evaluation Journal-urged practitioners to come forward and present their ways of dealing with knotty issues that come up in their evaluation practice. His reasoning was that there were many difficulties in evaluation theory that practitioners successfully engaged with and resolved without necessarily posing them or their resolution in theoretical terms and that if such practitioners did come forward, their practical knowledge could be mined and shared. I have now been a practising evaluator for more than fifteen years and the calls of these authorities embolden me to come forward and present a practitioner's viewpoint as well as some the ideas I use to inspire myself when I look about me and see the apparently waning fortunes of the profession in the Australian Public Service where I now work.

I have seen a profession in decline. In the 1980's the South Australian Public Service was a leader in introducing managerialism. At one of the many reorganisations of the SA Education Department¹, a Public Service Commissioner, Commissioner Cox (who incidentally regarded himself as an evaluator) told the assembled professional Education Officers that they had to accept a new principle-that of management being over the professions. I remember being amazed that the bald claim went unchallenged and it was years later before I realised other things also puzzled me about it. If management was above the professions then obviously it could not itself be a profession. How then was management to be understood and on what basis was the claim made? Well, perhaps the authority comes because management is serving some lofty good purpose. In that case, for practical purposes, I think that life itself can be taken as a paramount lofty purpose and since we are discussing human-to-human issues, authentic human living can be such a lofty good purpose.

The phrase 'authentic human living' was used by a philosopher who made a life's work out of studying human knowing. Lonergan (1967) saw human knowing as a dynamic structure necessarily involving the operation of sensing or experiencing (seeing, hearing, smelling, touching, tasting), of understanding (also involving inquiring, imagining, conceiving) and judging (also involving reflecting, weighing the evidence). Experience and understanding cannot, neither singly nor together, ensure that a person can come to know something. If knowledge were is only a combination of experience and understanding then we are easily deluded since there is no critical reflection as to whether or not our understanding is correct. As soon as we achieve an understanding of an experience, we would act on it without critically evaluating whether or not it is the understanding of our experience. Judgement is needed. We use judgement to distinguish fact and fiction, history and legend, chemistry and alchemy, astrology and astronomy. Once a person validly knows something (ie. has made judgements about his or her understanding of experience), then, if we follow Lonergan and I propose we should, then she or he has a basis for authentic human living. Since the fullest human life involves the judgements of evaluation, the evaluator has good grounds to reject Commissioner Cox's managerialist doctrine and say instead 'No. We are partners and we share your rights as manager'. What then, is the related responsibility to this assertion of rights? I think it is that evaluators must strive to do true evaluations.

According to the late Professor Freeman of UCLA, the idea of evaluating social action programs can be attributed to a sociologist in an American University in the 1930's. It was a time when the social sciences had enormous self-confidence associated with the popularity and insights of the so-called New Social Research which included the philosophy of the Vienna

Circle, psychiatry, psychology and sociology. It was also a time when President Franklin Roosevelt was delivering his controversial 'new deal' social action programs. This sociologistⁱⁱ made the claim that the New Social Research made it possible to decide how good the programs were. That is, that the programs could be evaluated. The evaluator's contribution was to carefully examine the achievements of the programs and applying the knowledge then emerging in the universities, pass judgement on their worth. I suppose the professional elements or responsibilities befalling the evaluator were for the comprehensiveness, reliability and defensibility of the judgement passed. It was not an undertaking to develop proposals for improvement.

It is not necessary for an evaluation to produce recommendations to be a good or true evaluation and recommendations do not automatically materialise just because an evaluation has been done. The reverse holds as well. It is perfectly possible to propose recommendations that do not come from an evaluation. This mutual non-implication gives rise to two dangers for an evaluator who is expected to produce recommendations:

1. That in developing and justifying the recommendations the evaluator fails to conduct a true evaluation
2. That in concentrating on doing a true evaluation the evaluator expeditiously adds recommendations that turn out to disappoint those who are supposed to act on them.

I will report what I was taught as an apprentice evaluator to deal with these dangers but first I should dwell on what counts as a 'true evaluation'. Tony Shinkfield and Dan Stufflebeam (1985) explored this question extensively in their book 'Systematic Evaluation'. They maintained that if a study does not report how good or bad something is, to all key stakeholders, it is not evaluation. Different values may be involved and conclusions may be contingent upon the value orientation adopted but nevertheless, value judgements must be madeⁱⁱⁱ.

Studies that are subordinated to political or expedient considerations and which are not released to the stakeholders or which are only partially released, are labelled as pseudo evaluations by Stufflebeam and Shinkfield. Pseudo evaluations include those studies in which the results are just made up rather than produced by systematic investigation. Studies they label as quasi evaluations include those that only look at the extent to which objectives have been achieved and scientific studies in the mould of pure research. Quasi evaluations can provide information that is useful and may help subsequent judgements to be made but fail to qualify as true evaluations because they do not expressly inquire into the worth or merits of the objectives or the scientific information produced. In the set of quasi evaluations I include studies that check the implementation of some management model (eg. Program Management and Budgeting in the Public Service) without examining the worth or merit of the model and its applicability.

I know that some freelance consultants say that the paying market is insignificant for studies in line with pure evaluation theory and that the selling point of an evaluation is its utility value in making improvements. Some of these also say that the text book world of evaluation is seriously flawed. Frankly I doubt both opinions and believe that it is poor marketing if we try to sell evaluation for its utility value. I have already proposed one reason for committing to true evaluation-the centrality of evaluation in coming to know things and for authentic human living. Another reason is another facet of the same gem, that when something is to be evaluated it is not some obscure technical variation that is wanted. For example, when it is publicly announced that something is to be evaluated, I believe that the citizen expects that there will be an impartial, objective assessment of that thing's worth or merit. And it is very much for this reason that the Joint Committee on Standards for Educational Evaluation set about to identify principles that when addressed should result in evaluations that withstand criticism and prove reliable.

The Joint Committee produced a comprehensive set of standards for educational evaluation in 1981, which became a defacto reference for program evaluation standards in fields other than education-so much so that when they revised the standards in

1994 they were emboldened to present them as 'The Program Evaluation Standards'. I have presented a summary of them as appendix 1. In the Joint Committee's view, evaluation is the systematic assessment of the worth or merit of some object. The 'object' need not be a program. It could be an institution or a product such as a policy.

Returning to the two dangers of producing recommendations. In both cases the evaluation standards are relevant and helpful especially in discouraging the second listed danger. There is another safe guard that is consistent with doing a true evaluation. I was taught that the evaluator had an obligation to be available to assist in the use made of an evaluation but not responsibility for them. If the evaluator insists on having two phases in an evaluation the risks of both dangers are reduced. The two phases are:

1. producing the findings that met the information needs of the stakeholders and reporting on the value or merit of the thing being evaluated- for which the evaluator is responsible; and
2. using the findings report, plan changes or develop recommendations-a phase for which the evaluator is not responsible but on which working party the evaluator should participate.

A strength of this approach is that the evaluation is passed over to a body specially set up to act on the evaluation findings and which can include those responsible for making operational or policy decisions. The evaluator on this separate body serves as a guide to the knowledge produced by the evaluation, makes sure that the full weight of respective findings is sufficiently appreciated and that related recommendations do justice to the issues insofar as they appeared to the evaluation team.

Conceivably, the two phases do not have to be sequential in time. As such, they are not unusual in a formative evaluation. Even if the manager of whatever is being evaluated or if the evaluator's boss are reluctant to set up two formally separate bodies the evaluator can usually arrange for two phases to happen. I do this by producing a Findings Draft Report. This is then sent to the manager of whatever is being evaluated inviting him or her to suggest recommendations in view of the findings reported. In practice, the manager rarely takes the lead in proposing recommendations and so, to encourage input, I begin drafting what we call an Interim Draft which has a recommendations section structured as follows:

2.1 Finding

(a summary statement of a finding)

2.2 Current Initiatives

(a description of what the manager or operational staff are currently planning or doing to improve matters in that finding area as understood by the evaluation team)

2.3 Further recommendation

(if the current initiatives appear likely to improve matters sufficiently well then this may be simply to state that the evaluation has 'no further recommendation'. Otherwise, a further recommendation is proposed with as much input from knowledgeable people NOT on the evaluation team as possible but in the light of the evaluation finding-subject to permission about confidential material being released in the process.)

2.1 Finding

(etc)

2.2 Current Initiatives

(etc)

2.3 Further recommendation

(etc)

This approach provides a framework for getting the kind of input from decision-makers that helps to make recommendations relevant and useful. Incidentally, the Interim Draft becomes the Final Report after the manager of whatever is being evaluated is given opportunity to formally respond to the Interim Draft. That response (if any) is included in the report with rejoinder (if any) by the evaluation team. This procedure of getting a response is copied from the internal auditors and has the tendency to encourage the implementation of the recommendations.

It is wise to prepare the way for how recommendations will be produced when determining the terms of reference of the evaluation. In the 1970's Joseph Wholey noted that it was difficult, sometimes impossible, to undertake evaluations of public programs because managers and other stakeholders resisted, were uncooperative or failed to grasp the purposes of studies. He also found that as a rule evaluation results were not used to refine and modify programs. When evaluations were completed, government policy makers and managers usually found that evaluations were too slow, inconclusive or answered the wrong questions. Wholey concluded that a systematic examination should precede any typical evaluation to make sure that a program is ready for evaluation and to identify the parts that can be evaluated. I think that the place of recommendations should also be included in that systematic examination.

Wholey called this preparation the pre-evaluation and I have found it a very useful insurance policy that guarantees that the evaluation is at least credible if not actually successful. I have no qualms proposing that up to 25% of the available resource days might be invested in a pre-evaluation study and report^{iv}. As an appendix, I include things that I believe should go into a pre-evaluation. I think it is wise to be continually planning how to produce and defend recommendations starting with the pre-evaluation.

Recommendations based on an evaluation should share the rigours expected of evaluation. These include an expectation that the various stakeholder perspectives and values deserve respect and should not be arbitrarily overlooked and some safeguards against the importation of concepts or paradigms that unjustifiably serve one interest group more than others. I think that this is one of the things that gives an evaluation its status—balance.

The evaluation standards (eg. U1, U2, U3, U4, P1, P5, P6, F2, A2, A3, A4, ...) are pertinent. I also believe that a principle that Dr Barry Shaw stressed at one of the monthly meetings of the ACT Branch of the Evaluation Society earlier this year is crucial: '(when doing evaluations) above all, understand and respect the system, its people, its shortcomings and value its strengths'.

My own philosophy on recommendations also applies Lonergan's theory of knowledge. I believe that producing recommendations is a matter of nominating some judgements or concepts (understanding) for preferment because judgements and concepts are the keys to controlling future experiences and reinterpreting past experiences. There is a related responsibility-to use the privilege in good faith in the best interests of those for whom the work is being done and with regard to other social and professional obligations. This is why the evaluator who values true evaluation is such a suitable agent to do this work. Such an evaluator is less likely to implant some currently-in-vogue management theory without checking to see that the experiences that are likely to follow are in accord with prior values, otherwise, and making sure that those affected are informed accordingly. The default driver in producing recommendations should, I think, be the next set of experiences that those affected by the recommendations either as actors or affected subjects want.

Paula Skippon, in her review of the Program Evaluation and Review Function of the Inspector-General Division of the Australian Department of Defence, noted that the issue attracting the most negative comments was the quality of recommendations. She applied four tests on the quality of recommendations: They should not be too numerous, they should be prioritised, well substantiated and well formulated.

Well-substantiated recommendations are more likely if the evaluation design is chosen so that the areas investigated match or imply the areas where recommendations are needed. For example, time series analysis and econometric analysis is unlikely to be a suitable evaluation design to provide a basis for recommendations to improve the quality of outputs. For such recommendations to be well based, the design probably should include examination of processes by someone competent to judge the value of alternative processes.

One performance indicator that is sometimes used to judge the success of an evaluation is the extent to which recommendations are accepted and acted upon. The case for such an indicator is a statistical dependency. Recommendations must be persuasive and realistic, relevant (apt) and timely, beneficial, reasonable and responsible. These qualities are more likely to be evident in recommendations if the evaluation also has such qualities. So acceptance of a recommendation is face value evidence that the positive requirements have been met.

Even though the preparation of recommendations is not a necessary part of a true evaluation, I encourage evaluators to accept the extra responsibility, but advise some safeguards. One might be humility. It is a great privilege and responsibility to affect where effort and resources should go. It is not something that happens easily either. An evaluation report and its recommendations are not likely to be the only basis of action and the evaluator needs to understand this. All the canvassing, consulting and integration of stakeholder interests that the evaluation does is tapping into something that not only set things up and got them going before the evaluation came on the scene, but will continue to do so. The evaluation and its recommendations will most often be an indicator of currents or alternative directions rather than the creator of the currents or the force that carves a new channel.

One the other hand, when the agent that commissions an evaluation insists upon getting recommendations, it can be a warning sign that the evaluator is being asked to do work that is the responsibility of someone else. This can be exploitative. A safeguard is to carefully map out respective roles and responsibilities as part of a pre-evaluation study, and perhaps to nominate the question of responsibility for operational development as an area to be evaluated and for which recommendations are to be produced.

A more pragmatic warning is that the work of developing recommendations will consume resources that might be needed for the evaluation. Dan Stufflebeam probably understands this best. This is why he put so much store in developing the evaluation standards. For him they served two purposes:

1. to help the evaluator keep his or her feet in the heady world of working with decision makers—to help confront political reality and not to misuse power
2. to help resolve disputes about the adequacy of an evaluation.

Referring to standards is a helpful way of managing the various trade-offs that may be necessary to satisfy resource constraints and for knowing when a study is no longer sufficiently balanced or comprehensive to qualify as an evaluation as the term is generally understood. Noteworthy in Australia is the recommendation of the Australian National Audit Office, that agencies doing program evaluation should establish their own standards (which may not necessarily be the American Standards) (ANAO, 1992). I think we can conclude that a professional evaluator has the right to negotiate standards by which the results will be judged and to register an opinion on the limitations of a study which appears to fall short of being a true evaluation as the term is generally understood.

As already stressed, one way of developing recommendations is to involve the decision-makers or staff involved in the thing being evaluated. Sometimes special knowledge is needed. If specialist or operational staff come to be included in the evaluation team, the evaluator or evaluation team leader should make sure that such staff are fully briefed about their roles. Training workshops on how to sit through interviews without interjecting or putting respondents right are usually a necessary precaution. I recommend a written specification of things that specialist or operational staff will be involved in and things that they are not expected to be involved in. Such a document can be a helpful tool to manage their contribution and to contain it to the areas of their expertise. Lastly, where possible, the recommendations, once developed, should be promoted using hard sell techniques.

In conclusion, individual evaluators must be kind to themselves. They will be humiliated from time to time. They will find themselves participating in covert evaluations and quasi evaluations and their attempts at true evaluations and implementation of standards will be disappointing. However, we should not be discouraged or get too overwrought. If, as an internal evaluator we must do an evaluation that assumes some political or ideological framework or agenda, accept the bias rather than promote or dispute it and let the method and standards reveal it—if it really is a bias. The framework may be surprisingly well justified. What matters is our faith in the endeavour. Evaluation may not be the highest good on its own, but it is an essential part of it, which I see as that noble activity of authentic human living.

REFERENCES

ANAO 1992, 'Auditing Program Evaluation Standards', *ANAO Performance Auditing Guide*, Australian National Audit Office, Australian Government Publishing Service, Canberra.

Department of Defence 1996, *Program Evaluation Report Program Evaluation and Review Function of the Inspector-General Division*, Directorate of Publishing, Defence Centre, Canberra

Department of Finance and Australian Public Service Board 1987, *Evaluating Government Programs (Financial Management Improvement program)* Australian Government Publishing Service, Canberra

Doyle C. J. 1992--notes prepared for a health evaluation conference (Melbourne?)

Freeman H. C. c.1980,--review report of the state of evaluation for UNESCO (Paris?)

Loneragan Fth B. 1967, 'What is knowing' collection by Crowe F. E., Herder & Herder, N.Y., N.Y.

Shadish Jr. R., Cook T. D., Leviton L.C. 1991, *Foundations of program evaluation : theories of practice*, Sage Publications, Newbury Park, CA.

Stufflebeam D. L. and Shinkfield A. J. 1985, *Systematic Evaluation: a self instructional guide to theory and practice*,: Kluwer-Nijhoff, Boston

The Joint Committee on Standards for Educational Evaluation, Sanders J. R., Chair 1994, *The Program Evaluation Standards 2nd Edn: How to Assess Evaluations of Educational Programs*, Sage Publications, Thousand Oaks, CA.

ⁱ At about this time the newly badged Chief Executive Office had voluntarily relinquished or was trying to relinquish his statutory responsibility for deciding the curricula of schools in the interests of a purer Managerialism. He did not think that he should have the authority to nominate curricula that could commit the Government to a level of costs to which it may not agree.

ⁱⁱ My apologies for not being able to provide this gentleman's name. Professor Freeman's review report of evaluation for UNESCO in which he provided this account is kept in the SA reference library (c. 1980). When I am next in Adelaide I may look it up, if anyone wants to register his or her wish to know with me.

ⁱⁱⁱ Following Shadish, Cook & Leviton (1991, p.101), Professor House proposes that an evaluation can report as follows: assuming that X (some summary of the values drawn from different stakeholders or from prescriptive theories) is held to be important, then Y (a thing being evaluated) is good for the following reasons

^{iv} I am indebted to notes prepared by Dr Colleen Doyle in 1992 for this account of Professor Wholey's work. I have little idea how I got Dr Doyle's notes because, as I suspect is common among practitioners trained on-the-job as I am, many references come second hand. I think I acquired them when rifling through the papers of one of my bosses while I was acting in the job and occupying the boss's room.

Appendix 1

THE PROGRAM EVALUATION STANDARDS

Sound evaluations of educational programs, projects, and materials in a variety of settings should have four basic attributes:

- Utility
- Feasibility
- Propriety
- Accuracy

The Program Evaluation Standards, established by sixteen professional education associations, identify evaluation principles that when addressed should result in improved program evaluations containing the above four attributes.

Dr. James R. Sanders, Chair
The Joint Committee on Standards
for Educational Evaluation
The Evaluation Centre
Western Michigan University Kalamazoo,
Michigan 49008-5178
616-387-5895

Sage Publications, Inc.
2455 Teller Road
Thousand Oaks, CA 91320
805-499-0721

JCSEE PR-1994
Approved by the American
National Standards Institute
as an American National
Standard. Approval date:
March 15, 1994.

This tear out is not copyrighted material. Reproduction and dissemination are encouraged.

Utility

The utility standards are intended to ensure that an evaluation will serve the information needs of intended users.

- U1 Stakeholder Identification** Persons involved in or affected by the evaluation should be identified, so that their needs can be addressed.
- U2 Evaluator Credibility** The persons conducting the evaluation should be both trustworthy and competent to perform the evaluation, so that the evaluation findings achieve maximum credibility and acceptance.
- U3 Information Scope and Selection** Information collected should be broadly selected to address pertinent questions about the program and be responsive to the needs and interests of clients and other specified stakeholders.
- U4 Values Identification** The perspectives, procedures, and rationale used to interpret the findings should be carefully described, so that the bases for value judgements are clear.
- U5 Report Clarity** Evaluation reports should clearly describe the program being evaluated, including its context, and the purposes, procedures, and findings of the evaluation, so that essential information is provided and easily understood.
- U6 Report Timeliness and Dissemination** Significant interim findings and evaluation reports should be disseminated to intended users, so that they can be used in a timely fashion.
- U7 Evaluation Impact** Evaluations should be planned, conducted, and reported in ways that encourage follow-through by stakeholders, so that the likelihood that the evaluation will be used is increased.

Feasibility

The feasibility standards are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

- F1 Practical Procedures** The evaluation procedures should be practical, to keep disruption to a minimum while needed information is obtained.
- F2 Political Viability** The evaluation should be planned and conducted with anticipation of the different positions of various interest groups, so that their co-operation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted.
- F3 Cost Effectiveness** The evaluation should be efficient and produce information of sufficient value, so that the resources expended can be justified.

Propriety

The propriety standards are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

- P1 Service Orientation** Evaluations should be designed to assist organisations to address and effectively serve the needs of the full range of targeted participants.
- P2 Formal Agreements** Obligations of the formal parties to an evaluation (what is to be done, how, by whom, when) should be agreed to in writing, so that these parties are obligated to adhere to all conditions of the agreement or formally to renegotiate it.
- P3 Rights of Human Subjects** Evaluations should be designed and conducted to respect and protect the rights and welfare of human subjects.

- P4 Human Interactions** Evaluators should respect human dignity and worth in their interactions with other persons associated with an evaluation, so that participants are not threatened or harmed.
- PS Complete and Fair Assessment** The evaluation should be complete and fair in its examination and recording of strengths and weaknesses of the program being evaluated, so that strengths can be built upon and problem areas addressed.
- P6 Disclosure of Findings** The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation, and any others with expressed legal rights to receive the results.
- P7 Conflict of Interest** Conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results.
- P8 Fiscal Responsibility** The evaluator's allocation and expenditure of resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible, so that expenditures are accounted for and appropriate.

Accuracy

The accuracy standards are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth of merit of the program being evaluated.

- A1 Program Documentation** The program being evaluated should be described and documented clearly and accurately, so that the program is clearly identified.
- A2 Context Analysis** The context in which the program exists should be examined in enough detail, so that its likely influences on the program can be identified.
- A3 Described Purposes and Procedures** The purposes and procedures of the evaluation should be monitored and described in enough detail, so that they can be identified and assessed.
- A4 Defensible Information Sources** The sources of information used in a program evaluation should be described in enough detail, so that the adequacy of the information can be assessed.
- A5 Valid Information** The information gathering procedures should be chosen or developed and then implemented so that they will assure that the interpretation arrived at is valid for the intended use.
- A6 Reliable Information** The information gathering procedures should be chosen or developed and then implemented so that they will assure that the information obtained is sufficiently reliable for the intended use.
- A7 Systematic Information** The information collected, processed, and reported in an evaluation should be systematically reviewed and any errors found should be corrected.
- A8 Analysis of Quantitative Information** Quantitative information in an evaluation should be appropriately and systematically analysed so that evaluation questions are effectively answered.
- A9 Analysis of Qualitative Information** Qualitative information in an evaluation should be appropriately and systematically analysed so that evaluation questions are effectively answered.
- A10 justified Conclusions** The conclusions reached in an evaluation should be explicitly justified, so that stakeholders can assess them.
- All Impartial Reporting** Reporting procedures should guard against distortion caused by personal feelings and biases of any party to the evaluation, so that evaluation reports fairly reflect the evaluation findings.
- A12 Metaevaluation** The evaluation itself should be formatively and summatively evaluated against these and other pertinent standards, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses.

Guidelines and illustrative cases to assist evaluation participants in meeting each of these standards are provided in *The Program Evaluation Standards* (Sage, 1994). The illustrative cases are based in a variety of educational settings that include schools, universities, medical and health care fields, the military, business and industry, the government, and law.

APPENDIX 2

PRE-EVALUATION

Things to be considered in a pre-evaluation include the following (not all items need to be formally reported on):

- a description of the thing to be evaluated
- identification of stakeholders and circulation lists for distribution of the report
- consultation on the information needs of stakeholders, researching and suggesting key evaluation issues, evaluative criteria and identifying areas where recommendations are expected
- confirming the main reason or purpose of the evaluation and advising stakeholders of it
- proposing a design for the evaluation that will include getting information likely to be necessary for producing and/or defending recommendations, estimating timeliness (usually using a project management package like Microsoft Project), checking the adequacy of the skills and credibility of the evaluation team and estimating resource requirements including costs
- selecting and documenting applicable technical standards
- recommending on the membership of the evaluation team
- arranging involvement of technical consultants (including program staff), planning training for them and drafting written guidelines for their participation
- recommending on whether a steering committee is desirable and, if so, nominating its members and confirming its charter (eg. will it be an advisory body or will it own the evaluation report and clear it for release? Will it undertake responsibility for developing recommendations (phase 2) or for acting on the evaluation?)