DOCUMENT RESUME

ED 202 869   TM 810 286

AUTHOR   Kemmis, Stephen
TITLE   Seven Principles for Program Evaluation in Curriculum
              Development and Innovation.
PUB DATE   Jan 80
NOTE   34p.

EDRS PRICE   MF01/PC02 Plus Postage.
DESCRIPTORS   Accountability; *Curriculum Development;
              *Definitions; Evaluation Methods; Foreign Countries;
              *Program Evaluation; Program Improvement
IDENTIFIERS   Australia

ABSTRACT
         The definition of evaluation as a process of
delineating, obtaining and providing information useful for making
decisions and judgments about educational programs and curricula
highlights the function of evaluative information of assisting
decision-making. It reflects the concept of the role of evaluation in
informing action at discrete decision points. It is important to
recognize that a curriculum program and its evaluation are highly
interactive throughout the curriculum development process. It is
critical in deciding upon a definition to guide evaluation efforts
that due importance is given to the pervasiveness of the evaluative
dimension of all human activity, and to the fact that it is present
in a range of individual and public judgment processes which exist
whether or not an evaluation is formally commissioned or expected of
project and program participants. Seven principles are offered which
attempt to emphasize that evaluation forms a natural part of the
critical thinking that guides the development process. (Author/GK)

SEVEN PRINCIPLES FOR PROGRAM EVALUATION

IN CURRICULUM DEVELOPMENT AND INNOVATION

Stephen Kemmis

Deakin University

January, 1980

SEVEN PRINCIPLES FOR PROGRAM EVALUATION

IN CURRICULUM DEVELOPMENT AND INNOVATION

Stephen Kemmis *

DEAKIN UNIVERSITY

ABSTRACT

Definition: Evaluation is the process of marshalling information and
arguments which enable interested individuals and groups to participate in
the critical debate about a specific program.

The principles:

1.   The principle of rationality as reasonableness

Program participants act reasonably in the light of their circumstances
and opportunities. It is the task of an evaluation to illuminate the
reasoning which guides program development and evolution, to identify
the contextual and historical factors which influence it, and to facilitate
critical examination of these matters in and around the program community.

2   The principle of autonomy and responsibility

Moral responsibility for an outcome can only be ascribed to a person
to the degree that his or her free choice of action as an autonomous
moral agent was a cause of that outcome. Curriculum development
projects and programs are cooperative enterprises. Evaluators must
illuminate the interactive character of accountability for a program.

3   The principle of community self-interest

When a curriculum development project is formed, it is a community of
self-interests -- it represents the self-interests of all participants
within its terms of reference. The evaluator has a responsibility to
illuminate the extent of commonality and conflict among the values
and interests of participants in this community.

---

4   <u>The principle of plurality of value-perspectives</u>
A range of different value-perspectives becomes relevant in judging
a program. An evaluation should identify these different perspectives
and be responsive to the different concerns they imply.

5   <u>The principle of the self-critical community: internal evaluation,</u>
<u>evaluation consultancy, meta-evaluation, external and independent</u>
<u>evaluation</u>
Critical debate about the nature and worth of a program already exists
within and around its program community.   It is the task of program
evaluation to refine this debate and improve its bearing on program
action.   Evaluation consultancy may provide additional tools for this
purpose. Meta-evaluation efforts may help to improve the quality of
the contribution of a program evaluation. An external evaluation may
contribute to the critical debate by increasing awareness of a particular
set of values and interests relevant to a program; it should not be
thought of as an alternative to the self-critical process.

An independent evaluation may help to harness
program self-criticism where the program community is diffuse or divided
by controversy.   Self-criticism by the program community is the
primary basis for program evaluation; other evaluation efforts extend
it in different ways but do not supplant it.

6   <u>The principle of propriety in the production and distribution of</u>
<u>information</u>
Evaluation processes inevitably affect the political economy of information
in a program (the production and distribution of information about it).
Because information and arguments justify or legitimise decisions,
evaluation affects the distribution of power and resources in program
situations.   Program participants and interested observers live with
the consequences of use and abuse of evaluation information.   An
evaluation should have explicit principles of procedure which govern
its conduct and its processes of information production and distribution.

7   <u>The principle of appropriateness</u>

Evaluation design is a practical matter. An evaluation must be
appropriate to the program setting, responsive to program issues, and
relevant to the program community and interested observers. An
evaluation design must be renegotiated as the study progresses in the
light of changing circumstances, issues and interests, and in the light
of its own consequences (as they become apparent).

Different definitions of evaluation abound. The Australian Curriculum Development Centre (CDC) Study Group on curriculum evaluation reviewed a variety of definitions with currency in the evaluation literature and adopted the following one as the most useful guide for the evaluation of CDC's own projects and programs and for curriculum evaluation more generally:

> *Evaluation is the process of delineating, obtaining and providing information useful for making decisions and judgments about educational programs and curricula.*[1]

This definition highlights the function of evaluative information of assisting decision-making. It reflects a fairly widespread agreement among evaluation theorists[2] about the role of evaluation in informing action at discrete decision-points. But it is important to recognise that a curriculum program and its evaluation are highly interactive, not only in "summative" decisions, but throughout the process of curriculum development. In short, the discrete decision points are few and far between, and evaluation permeates development: the two processes are not discontinuous. On the contrary, they are inseparable. Accordingly, a desirable definition of evaluation will acknowledge the mutuality of the relationship between evaluation and curriculum development and its continuous, organic and reflexive contribution to thought and action about a curriculum. Other definitions of evaluation, while more general than the one adopted by the CDC Study Group seem more likely to recognise the pervasiveness of the evaluation function, and less likely to treat it as discontinuous and separate from development. Stake and Denny, for example, have this to say:

> *Considered broadly, evaluation is the descovery of the nature and worth of something. In relation to education, we may evaluate students, teachers, curriculums, administrators, systems, programs and nations. The purposes for evaluation may be many, but always evaluation attempts to describe something and to indicate its perceived merits and shortcomings ... Evaluation is not a search for cause and effect, an inventory of present status, or a prediction of future success*

---

1   Curriculum Evaluation: A CDC Study Group Report, CDC professional Series, Curriculum Development Centre, Canberra, 1977, p.24

2   See, for example, Cronbach, L.J. "Course Improvement Through Evaluation", Teachers' College Record. 1963, 64, 672-693; MacDonald, B. "Briefing Decision-Makers.", in E.R.House (ed.) School Evaluation: The Politics and Process, Berkeley, Calif.: McCutchan, 1973; Stufflebeam, D.L. et al., Educational Evaluation and Decision Making in Education, Itasca, Illinois: Peacock, 1971.

> *It is something of all of these but only as they contribute*
> *to understanding substance, function and worth.*[3]

This definition, emphasising "nature and worth", "perceived merits and short-comings" and "understanding", goes further towards acknowledging evaluation as an ever-present aspect of considered action. One might quibble over the use of the word "discovery", suggesting as it does that the nature and worth of the thing evaluated antedate the evaluative search for them, but the force of the term is mitigated by the notion that merit and shortcomings are "perceived" rather than intrinsic or immanent.

It is critical in deciding upon a definition to guide evaluation efforts to give due importance to the pervasiveness of the evaluative dimension of all human activity, and to the fact that it is present in a range of individual and public judgment processes which exist whether or not an evaluation is formally commissioned or expected of project and program participants. Indeed, when evaluations of particular programs are commissioned, they should approximate (and focus and sharpen) these informal critical processes, not ignore or supplant them. Though the formality of commissioning or requiring an evaluation imposes certain obligations to formalise and discipline the individual and public judgment processes which occur naturally in considered activity, formal evaluations should attempt deliberately to preserve something of the conviviality of the informal processes.

Preserving conviviality is no easy task. These principles attempt to provide a framework within which conviviality can be preserved by emphasising the continuity and mutuality of concern between program participants, a program sponsor, an evaluation sponsor and an evaluator. They also attempt to emphasise that evaluation forms a natural part of the critical thinking that guides the development process. This is not to say that formal evaluations can lack rigour, discipline or honesty; rather, it is to assert that their critical edge should be tempered with humane values rather than narrowly technocratic or bureaucratic concerns.

Accordingly, the definition of evaluation which has informed and guided the development of these principles is this:

---

3   Stake, R.E. & Denny, T. "Needed Concepts and Techniques for Utilizing
     More Fully the Potential of Evaluation", in National Society for the
     Study of Education Yearbook LXVIII, Pt.II, Chicago: University of
     Chicago Press, 1969, p.370.

> *Evaluation is the process of marshalling information*
> *and arguments which enable interested individuals*
> *and groups to participate in the critical debate*
> *about a specific program.*

So construed, evaluation consists in harnessing and refining the
ubiquitous processes of individual and public judgment, not in resolving
or replacing them with a technology of judgment.

There is a certain kind of seduction in discussing principles: one
is often inclined to share the aspirations they embody simply because they
seem worthy ones or because the rhetoric of principles is lofty and stirring.
For this reason, the discussion of each principle includes a reference to
two alternative principles. Each of the three resulting sets of principles
(the set advocated here and the two alternative sets) are more or less self-
consistent.[4]

The reader is invited to choose between the three sets and to consider
in choosing that the choice may be a revealing one: it will indicate a
preference for the interests and concerns of one group over another when
evaluator, sponsor and program participants interact in the evaluation process.

1   THE PRINCIPLE OF RATIONALITY AS REASONABLENESS

> *Program participants act reasonably in the light of*
> *their circumstances and opportunities. It is the*
> *task of an evaluation to illuminate the reasoning which*
> *guides program development and evolution, to identify*
> *the contextual and historical factors which influence*
> *it, and to facilitate critical examination of these*
> *matters in and around the program community.*

Evaluation is always guided by the impulse to understand and to act on
the basis of understanding.   It thus has a major role to play in
articulating justifications of action. Properly speaking, the justification
of action is not merely a backward-looking enterprise, to be equated with
post-rationalisation. On the contrary, it is concerned with demonstrating
both how things have come to be as they are (that is, with illuminating the

---

4   The two "alternative" sets may be somewhat less internally-consistent
     since they parallel the first set.

reasoning which has guided the activities of those associated with a program and identifying the circumstances which shaped and constrained them) and with providing information and arguments which can justify contemplated action.

In evaluating an educational program it is critical to explicate the reasoning which has guided the activities of those associated with it. Unless there are very good reasons for assuming otherwise, the evaluator and the evaluation sponsor should assume that persons will in general act reasonably in the light of their current circumstances and the available opportunities. That is to say, it should be assumed that those associated with a program are committed to acting with understanding; in ways informed by their values and beliefs; wisely and prudently. By explicating the reasoning of those in and around a program, an evaluation may therefore share the understandings of those who are deeply concerned with it. It may help to disclose the nature of the program and the values it embraces and give those associated with it an opportunity to have their perspectives on it represented.

The truths to be told about educational programs are social truths. They are negotiated among those who claim to know it and those who want to know it better. The principle of "rationality as reasonableness"[5] draws attention to this negotiation process. Claims about the program are defended and challenged in a process of critical debate or conversation. What will count as the truths to be told about the program will depend upon the quality of the debate. Rational debate consists in giving reasons and defending reasoning with information and arguments.

An evaluation may make a substantial contribution to this critical process. It can gather evidence relevant to program aims and claims and subject it to critical cross-examination. It can elicit, articulate and share understandings about why the course of development, implementation and dissemination is as it is by reference to the purposes of participants, the constraints of circumstance and the available opportunities, and reflect on these understandings in the light of the wider context and experience of the program as a whole. Likewise, it can subject the views of program audiences -- supporters, detractors and those who have not yet a basis for

---

5   See Weir, E. "Rationality as reasonableness". Office for Instructional
     Research, University of Illinois at Urbana-Champaign, 1976

making judgments about it -- to critical scrutiny so a "conversation"
between the perspectives of those associated with it can be created and
maintained.

The implication of the principle of rationality of reasonableness is
that evaluators will attend to a wide variety of perspectives on a program,
to the diverse claims made about it, to its context and to its history.
They will thus be in a position to harness and refine the individual and
public judgment processes by which the program comes to be understood and
by which its value is determined. The quality of the evaluation may
be judged by the quality of its contribution to informing and improving
the critical debate about the program.

Evaluation should thus aim to contribute to program improvement both
directly and indirectly: by its direct interaction with program participants
and by feeding and refining the interaction between program participants
and their audiences.

An alternative principle to the principle of rationality as reasonableness,
and one which is not advocated here, is that of "rationality as rule-following".
Stake's label of "preordinate"[6] fits evaluation approaches which have prior
rules for judging a program and which do not respond to immediate value-
perspectives, information-needs and circumstances. For example, some
evaluation approaches are based on the notion of rational consumption and
set out criteria and standards which must be met before a program can be
considered a "good buy". Scriven's "product evaluation checklist"[7] is
perhaps the best example of this. It lists thirteen considerations in the
evaluation of products, producers and proposals, and sets standards of
adequacy for each. They are:

    1. need,   2. market   3. performance: truefield trials
    4. performance: true consumer
    5. performance: critical comparisons,
    6. performance: long term,
    7. performance: side-effects,

---

6   Stake, R.E., "To Evaluate an Arts Program". In R.E.Stake (ed.) Evaluating
    the Arts in Education: A Responsive Approach. Columbus, Ohio:Charles E.
    Merrill, 1975.

7   Scriven, M."Evaluation Perspectives and Procedures", in W.J.Popham (ed.)
    Evaluation in Education: Current Applications. Berkeley, Calif.: McCutchan,
    1974.

8. performance: process,
9. performance: causation,
10. performance: statistical significance,
11. performance: educational significance,
12. cost-effectiveness, and
13. extended support.

These are powerful considerations and the model provides a useful set of questions to be asked of a program or product. But the criteria are subject to interpretation in application and they do not respond sufficiently to the nature of the critical debate which actually attends a program. How would such an evaluation model account for the controversy over SEMP (the Curriculum Development Centre's Social Education Materials Project), for example, where different value premises underlie the opposing positions and they cannot be resolved on the basis of a common criterion test? Furthermore, the criteria are extremely stringent in practice and few curriculum developments are able to justify themselves in terms of the standards Scriven sets. Rhetorically, one might claim that these standards should always be met, but since they cannot be in real curriculum development how can what actually happens be justified? It must be by appeal to something else, that is, by something other than the criteria the checklist proposes.

There are other forms of rationality as rule-following. Evaluations guided only by considerations of testing and measurement are rule-following, drawing their rules from psychometrics. Cronbach[8] has pointed out their limitations for real curriculum evaluation which depends on more information than data about student performance on carefully-designed tests. More generally, Hastings[9] has drawn attention to the fact that research methods, inspired by their own rational models (for their own purposes) often distort evaluation problems by twisting them to fit the methods they employ. Clearly, methods which are designed to handle closely-defined, special problems are of limited help in handling poorly-defined, multiform program realities. The problem about specifying the rules for rational justification in curriculum is that the rules are likely to be limited; rationality as reasonableness makes no such prescription about what particular rules must be applied, though it recognises the usefulness of such rules as far as they seem relevant.

---

8   Cronbach, L.J., op.cit.

9   Hastings, J.T. "The Kith and Kin of Educational Measurers", Journal of Educational Measurement, 1969, 6, 127-130

A second alternative principle, also not advocated here, is the view of rationality implied in "rational planning". This view sees justification as based on the notion of satisfaction of needs. If a need can be identified and regarded as an urgent one, then programs can be designed to satisfy it. Relatively few educational programs can be said to satisfy urgent needs, though education as a whole responds to a general social need. But the rational planning approach tends also to take a contractual view of programs: to see them in terms of the obligations imposed on those brave or foolish enough to accept grants to develop programs. Within such a view, measurements of need-reduction or aims-achievement, coupled with fulfillment of contractual obligations, are sufficient to demonstrate that a program is justified.

Potential grantees exploit the invitation this approach suggests, "manufacturing" needs, overpromising, and using limited or biased measures. The exploitations are not always deliberate; rather, they are inspired by a cultural tendency towards legalism and concepts of exchange rooted in economics. Such values have their place, of course, and program evaluations which do not attend to the contractual obligations of grantees may fail to take account of important aspects of the programs.

But the "rational planning" approach to evaluation may treat educational programs in a bureaucratised way which does not do justice to the organic and reflexive character of social and educational life. Program objectives change, as they should, in response to changing circumstances and opportunities; educational programs rarely specify the sole means by which goals can be achieved; "needs" are usually relative in education, not absolute. The "rationality as reasonableness" approach is likely to take a more open-minded view of program justification which is sensitive to the relativity of educational values and their adaptation to social contexts.

To adopt the principle of rationality as reasonableness is thus to take the view that social truths are socially-negotiated and historically- and culturally-relative. It is to reject the notion that any discrete set of rules can be formulated which will provide universal criteria of program adequacy. Similarly, it is to reject the notion that programs can be justified solely by reference to their own goals, objectives and obligations. or by reference solely to needs-reduction. Evaluations based on either of

12

these alternative perspectives are likely to be limited and partial, providing an inadequate basis of information and argument for those who want to enter the critical debate about a program.

2   THE PRINCIPLE OF AUTONOMY AND RESPONSIBILITY

> *Moral responsibility for an outcome can only be ascribed to a person to the degree that his or her free choice of action as an autonomous moral agent was a cause of that outcome. Curriculum development projects and programs are cooperative enterprises. Evaluators must illuminate the interactive character of accountability for a program.*

Just as the evaluator may assume that those involved with a program act rationally in the sense that they are open to arguments based on reason, so it may be assumed that those involved with the program are autonomous and responsible moral agents. This has implications for the way program participants, evaluators, evaluation sponsors and program sponsors view accountability issues in evaluation.

Most program sponsors use or distribute public funds for program development and implementation. They are publicly accountable for their use of these funds; program participants must also account for their use of the resources allocated to them. Financial and management procedures usually accompany development project fundings to ensure that accountability demands are met. A "maximalist" view of accountability requires program participants to justify every decision about the use of resources by reference to program goals, social needs and the consequences of each decision (especially i terms of program outcomes for students and teachers). But it is sufficient to adopt a "minimalist" view of accountability as keeping financial and other records which show that programs have operated within their budgets and according to their terms of reference, and to make these records open to view.[10]

---

10   See Stake, R.E. "School Accountability Laws", Evaluation Comment, 1973, 4, 1-3. The present definition is based upon Stake's.

More generally, the minimalist view of accountability is based on the
principle of autonomy and responsibility. According to this principle,
moral responsibility for an outcome can only be ascribed to a person to the
degree that his or her free choice of action as an autonomous moral agent
was a cause of that outcome. To the degree that the person's choices were
constrained by others or by circumstances outside his control, then to that
degree the person cannot be held responsible (or at least not solely
responsible) for the outcomes.

In a climate where the accountability dragon has reared its head and
begun to roar in education, it is as well to be clear about these issues.
Any program sponsor is always implicated in the accountability issue. By
constraining the choices open to those who carry out its work, it reduces
their autonomy as agents. It must therefore accept a part of the responsibility
for the outcomes. Accountability is always bilateral: it concerns provision
as well as performance.

Accountability is not a matter of the distribution of praise and blame.
Curriculum development is a cooperative enterprise between program sponsors
and participants; the co-actors share responsibility just as they jointly
constrain one anothers' opportunities. Nor will it do to take an overly-
personalised view of responsibility. Structural constraints of program
design and organisation impose constraints on free action whose effects are
sometimes difficult to predict; circumstances which surround development are
often beyond the control of participants and may restrict free action in
unanticipated ways. One task of program evaluation is to identify such
structural constraints and to determine their effects. A program evaluation
should therefore be highly sensitive to historical and contextual issues
so that the work of the program can be seen against its background of
constraints and opportunities.

One alternative principle concerning accountability not advocated here
would be one based on ideals of truth and justice. According to such a
principle, an evaluator or program sponsor might adopt some view of what
constituted true and just work, perhaps spelling out criteria for truth and
justice. These would then constitute a view of what "the good" (or best)
in curriculum development might be. Program participants could be held
accountable for deviations from this ideal.

14

This sort of principle is clearly unsuitable given the commitment already declared to the notion that social truths are socially-negotiated. And it is unsuitable in a pluralist society where different value-perspectives with different patterns of coherence and legitimacy coexist. Curriculum development always expresses social and educational values and it is proper that they be critically analysed and examined in each case. Far from asserting what values are proper for a program and then judging it according to those values alone, evaluations should attempt to explore the diverse values and value-perspectives expressed in a program and the work of those involved in developing it, setting these in a context of the diverse values of the wider society beyond the development group.

A principle concerning accountability based on ideals of truth and justice and concerned with attributing blame for deviations from these ideals thus seems both inappropriate and unworkable. The principle of autonomy and responsibility allows for interaction among value-perspectives; it does not close off critical debate by imposing an ideal because this happens to be the ideal of the evaluator or the program sponsor. Program sponsors will no doubt have their own curriculum values and hope to express them through their work, but they should not assume that these values can be imposed unilaterally -- they must stand up to the test of critical debate among a plurality of value-perspectives.

A second alternative view of accountability, likewise not advocated here, might be one based on the notion of contractual obligation. Such a view would seem to be based more on notions of prudence and expedience than on principle. Nevertheless, it is worth exploring briefly. The accountability issue as it has been aired in education has frequently been discussed in these terms. According to this view, there is a chain of obligations from the classroom teacher through education systems to Ministers of Education and ultimately through Parliaments to the people. Each superordinate agency is seen as totally responsible for the actions of all subordinates. This view is based on a notion of management which might be described as highly positivistic, that is, the notion that management causes events to occur. 1. is managerialism is contrary to the facts of development, of course: subordinates are not mere operators whose every action is determined by job specification. It is also contrary to the values of professionalism in education: teachers are not operatives but relatively autonomous professionals. (This value reaches its peak in the notion of academic freedom; it is moderated by notions of social responsibility).

If a program sponsor were to adopt this contractual and managerialistic view of accountability it might present itself as a "responsible" authority yet preserve the capacity to disavow responsibility when things go wrong, claiming that operatives in its projects and programs stepped outside their specified tasks and responsibilities. Naturally, it cannot do so: the principle of autonomy and responsibility embodies an acknowledgement of the interactive character of accountability, and the fact that it is bilateral or multilateral, not unilateral.

Ideals for action as a basis for accountability are thus inadequate as are purely contractual views. There is merit in both alternatives: the one puts a premium on the value-commitments of program sponsors, the other puts a premium on the responsibility of program staff to meet their contractual obligations. But each is insufficient, failing to recognise value-plurality and the cooperative character of development work (a program sponsor is not simply an initiator of development activity; it negotiates the character and amount of development activity in planning and executing its programs).

3   THE PRINCIPLE OF COMMUNITY SELF-INTEREST

> *When a curriculum development project is formed, it is a community of self-interests -- it represents the self-interests of all participants within its terms of reference. The evaluator has a responsibility illuminate the extent of commonality and conflict among the values and interests of participants in this community.*

When a program sponsor enters relations with other agencies (education systems or individual officers who work on its projects, for example) and negotiates the terms of reference which set up a project or program, it forms a community of self-interests with them. Within the terms of reference of the program, the (self)interests of the sponsor and these other individuals or groups coincide. Program organisation always has this cooperative character. A program sponsor is thus always only one among a number of participants in the communities of self-interests formed by its projects and programs.

Individual self-interests which exist outside the terms of reference of

the cooperative are irrelevant to it unless conflicts of interest prejudice the interests of the cooperative itself.

When a program sponsor commissions a project or sets up a program, it establishes a community of self-interests. Similarly, when a program sponsor commissions an evaluation study, it should recognise that the role of the sponsor is a critical aspect of the program to be evaluated. Sponsorship of development and evaluation confers no exclusive right to have the interests of the sponsor served at the expense of other participants in the community of self-interests. Both with respect to programs and their evaluations, the actions of program sponsors may be examined in terms of their fairness in agreeing terms of reference and negotiating contracts. House Care[11] set out conditions indicative of fair agreement, one of which is community self-interestedness. The conditions are :

> non-coercion
> rationality
> acceptance of terms
> joint agreement
> disinterestedness
> universality
> community self-interestedness
> equal and full information
> non-riskiness
> possibility
> "counting all votes"
> participation.

If these conditions attend the process of reaching the agreement, then it is a fair one, House and Care argue. Agreements so reached are binding on all parties and can only be overridden by appealing to some higher moral principle. The conditions for fair agreement represent an aspiration, of course: sometimes for "technical" reasons (for example, in relation to questions about who was able to attend a meeting where the organisation of a project was discussed, how negotiations proceeded over time, and who was involved at what stages) and sometimes for prudential reasons (for example, the need to unilaterally define the terms of reference for some activity in order to let contracts), these conditions will not be met.

---

11   House, E.R. & Care, N.S. "Fair Evaluation Agreement", Centre for
       Instructional Research and Curriculum Evaluation, University of Illinois
       at Urbana-Champaign, Urbana, Illinois 61801 (mimeo), 1977.

These conditions apply as much to program evaluation as to program development. In general and in accordance with the principle of community self-interests, the evaluation of a project or program should be regarded as a cooperative venture, not as an information service for a sponsor's own exclusive use. Sponsors should recognise that they will "co-own" the information generated by an evaluation study with other participants; and that their own role in shaping a program is relevant in evaluating the program.

There are respects in which program sponsors' evaluative efforts will not be entirely governed by this principle, however. Firstly, sponsoring agencies generally have obligations to collect financial and other information which allow them to discharge their statutory obligations; sponsors do this (one might say) in their own self-interest. This will not generally conflict with the interests of other participants and is simply a condition of sponsorship of developments. Secondly, sponsors initiate some forms of evaluative activity before a community of self-interests is formed. They do this as a necessary and private matter when they evaluate proposals for projects andprograms in the light of their own interests.

Once a community of self-interests has been formed, however, a program sponsor's particular interests must be considered in an evaluation study alongside the interests of other participants. A sponsoring agency cannot expect to withdraw from the cooperative enterprise at the point of judgment in the guise of "disinterested observer" (disavowing involvement). It is relevant to note that the concept of the interests of a sponsoring agency is a slippery one, at least in the case of government agencies, charitable foundations and the like. As institutions, these agencies are themselves communities of interests bound together by the common goals of their enterprise and the organisation of their common work. They may be defined by an Act, charter or constitution, governed by a Council, responsible to a Minister, and express a variety of interests in their staff. Within them, a variety of individuals bring their interests to bear in shaping the overall common interest. At the same time, these diverse individuals are capable of disinterest, suspending their own values and interests as they try to understand and develop the common work. The notion of a community of self-interests is an important one simply because it emphasises these "internal" and "external" negotiations. Programs are cooperative efforts among

participants and thus sponsors cannot disavow their involvement when
commissioning evaluation studies.

The foregoing paragraphs have emphasised the sponsor's involvement
in the cooperatives formed in project and program work primarily because
it should be clear both to project workers and to evaluators that the work
is cooperative. Naturally the same might be said of all other participants
in the cooperative: all are bound by it unless they withdraw for some
other overriding reason.

Two alternative principes to the principle of community self-interest
could be proposed, based on the one hand on the notion of "the public interest"
(defined outside the interests of participating individuals and agencies), or,
on the other, on the notion of the sponsor's self-interest. The first might
depend on some definition of what is supposed to be in the public good and
specify criteria by which programs might be judged; the second might assert
a narrow definition of value according to the sponsor's own perspectives.
For reasons already outlined in the discussion of earlier principles (social
negotiation of social truths, pluralism, the importance of contributing to
the critical debate about a program), neither of these approaches is a
reasonable option. Since the sponsoring agency is a participant in the
development process along with other groups and agencies, it is interested
in improving the quality of critical debate about curriculum. To do so,
it requires evaluations which share information among those involved in the
process (with due regard for the protection of the rights of individuals)
rather than evaluations which serve only its own purposes or only those of
other particular groups within the cooperative. (Equally, it is not
interested in evaluations which serve only the purposes of those outside
the communities of self-interests it forms).

4   THE PRINCIPLE OF PLURALITY OF VALUE-PERSPECTIVES

> *A range of different value-perspectives becomes relevant
> in judging a program. An evaluation should identify these
> different perspectives and be responsive to the different
> concerns they imply.*

Program participants' values and interests are served by their participation
in curriculum development. The particular individuals and agencies cooperating

in a project or program have their own values and interests which may be independently justified. Other audiences for the work of a project or program will likewise judge it by reference to their own values and interests. A program sponsor can claim no monopoly on the values or criteria by which a program is to be evaluated.

According to the principle of plurality of value perspectives, program evaluators should recognise that a range of perspectives may be relevant in making a judgment of a program. Any judgment of the value of a project or program will be made in the light of the value commitments of the judge; program evaluators should therefore inform themselves and their audiences about the value perspectives of relevant judges, be responsive to their concerns, and provide information which is appropriate (and valued most highly) as evidence according to their criteria of judgment. If the information and arguments collected in the course of the evaluative study are relevant and significant to the audiences of the evaluation, there is a greater likelihood that it will be used in the critical debate about the program.

As an alternative to this principle, the view could be taken that judgments of a program should be the prerogative of those best-equipped to judge: for example, competent authorities in curriculum as a field, curriculum development processes, the subject-matter of the program, or teaching and learning processes. While these specialists may well be able to provide valuable information and insights into the program, they are not the only ones entitled to judge it. An evaluation should embrace such perspectives but should go beyond them to take into account the perspectives of other interested parties (for example, students, parents, community groups or employers). The mature judgment of specialists may be of great value to audiences less familiar with specialist debates about the nature and worth of a program, but lay concerns demand attention too: as clients or observers of the program, laymen must have their questions treated seriously in an evaluation study, have specialist issues made accessible to them, and see how these specialist issues fit into the broader context of the issues concerning the program as a whole.

Still another principle which could be adopted would concern a sponsor's own right to judge, and the primacy of its right as a sponsor of development to have its own questions answered. To be sure, a program evaluation should

address questions which the sponsor regards as important. But such a principle, pursued single-mindedly, would have a conservative and defensive effect. It would make the evaluation a service for the sponsor at the expense of other audiences with legitimate rights to be heard. In order to feed the critical debate about the program and to refine it, an evaluation must engage the perspectives of a variety of audiences.

5   THE PRINCIPLE OF THE SELF-CRITICAL COMMUNITY: INTERNAL EVALUATION,
     EVALUATION CONSULTANCY, META-EVALUATION, EXTERNAL AND INDEPENDENT
     EVALUATION

> *Critical debate about the nature and worth of a program*
> *already exists within and around its program community.*
> *It is the task of program evaluation to refine this*
> *debate and improve its bearing on program action.*
> *Evaluation consultancy may provide additional tools for*
> *this purpose. Meta-evaluation efforts may help to*
> *improve the quality of the contribution of program*
> *evaluation. An external evaluation may contribute to*
> *the critical debate by increasing awareness of a*
> *particular set of values and interests relevant to a*
> *program; it should not be thought of as an alternative*
> *to the self-critical process. An independent evaluation*
> *may help to harness program self-criticism where the*
> *program community is diffuse or divided by controversy.*
> *Self-criticism by the program community is the primary*
> *basis for program evaluation; other evaluation efforts*
> *extend it in different ways but do not supplant it.*

The community of self-interest formed by a curriculum project or program is likely to embrace a variety of value-perspectives which, through their interaction in its life and work, create a continuing conversation about its nature and worth. This conversation provides a basis for systematic self-criticism within the community; it is nourished by contact with perspectives from the wider social and educational communities outside.

A major task for program evaluation is to harness this self-critical conversation: to collect the perspectives and judgments of those associated with a program, to reclaim meanings and concerns from the flux of program experience, and to make this store of understandings available to participants and other audiences. Describing the program, formulating issues regarded as significant by those associated with it, collecting judgments and portraying these in ways which are accessible to evaluation audiences -- these are activities

through which the evaluator can contribute to the critical debate about a program and improve the quality of the critique.

Such activities acknowledge that critical debate already exists independently of the evaluation of a program, within and between program staff and interested observers. Far from attempting to supersede "natural" critical debate, an evaluation should attempt to capitalise on it, harness it and refine it. In doing so, it can engage the perspectives and concerns of those within and outside the program (without imposing perspectives which are regarded as "foreign" by those within). By bringing these perspectives in contact with one another and opening up interaction across the borders of the program, the evaluation may generate authentic knowledge about the program -- that is, knowledge grounded in the life-circumstances of participants and understood as experience. Authentic knowledge is the only sure basis for program improvement and for improvement of critical debate because it reflects personal understandings which will express themselves as free commitments in the actions of participants.

The implication for evaluators is that evaluations should be responsive to audience concerns and the real, experienced issues which surround a program. The evaluation task is thus an educative one, informing and developing the understandings of those associated with the program. The evaluation may accept as its primary task the formulation program issues in ways which clarify them for program participants and audiences. It may report frequently rather than just towards the end of the evaluation, so that the perspectives of participants and audiences can be engaged more or less continually rather than in a single confrontation of perspectives. The recurring "reports" of the evaluator can be regarded as a conversation which develops the points of view of those it engages. In this conversational process, interim reports should be less formal and regarded as ephemeral (rather than highly authoritative) by participants and audiences. A final report should reflect the evolution or history of the critical debate in and around the program. Evaluation reporting should be regarded as a dynamic process rather than static or discontinuous; the evaluation findings as contributory and reflexive rather than controntational or inert; the evaluator as a facilitator of debate rather than as an "objective" outsider who represents truth against one-sidedness or complete understanding against the partial understandings of participants.

22

Regarding the program as a self-critical community does not mean that it is an insular group feeding only on its own perspectives; using the self-critical debate within the program as a basis for the evaluation does not mean that the evaluation becomes simply a kind of self-report. Through the evaluation (as well as through program initiatives), participants should be brought into contact with the perspectives of other relevant judges and audiences, some of whom may be quite distant from the program. The self-critical communityof the program can incorporate the perspectives of "outsiders" by creating a conversation with them through which both sides can learn each others' perspectives. This can occur if the program and the evaluation create opportunities for outsiders to see the work (or portrayals of the work), to consider it, to judge it, and to explain their judgments.

The principle of the self-critical community establishes self-criticism as the cornerstone of program evaluation. All participants in the community of self-interests formed by the program have a right to be heard in the critical process. As already indicated, the value of self-criticism does not preclude external judgment, rather, it attempts to create mechanisms whereby external judgments can be incorporated into project or program thinking. To emphasise the value of self-criticism is not to advocate program insularity; on the contrar y, it is to emphasise the value of authentic knowledge as a basis for development and debate and to encourage participants to take a broad, critical view of the program in its wider historical context. But it is also to stress that once a program surrenders self-knowledge to external authority as a basis for development, it loses its autonomy as an intellectual community.

As a corollary to this principle it follows that each participant agency in the cooperative enterprise of a program regards itself as a self-critical community, and evaluates its own activities in a spirit of self-criticism.

The primary implication of the principle of the self-critical community is that curriculum projects and programs should establish "internal" evaluation mechanisms which can systematically record and develop the critical debate about their work. These "internal" evaluations may be augmented in four ways: by evaluation consultancy, by meta-evaluation, by external evaluation, and by independent evaluation.

(a) Evaluation consultancy

Where specialist evaluation expertise is available, program participants
and evaluation sponsors may want to take advantage of it. Using such
advice is by no means precluded under the principle of the self-critical
community. For many evaluation tasks (like interview techniques,
questionnaire design, planning and sampling),technical assistance is
highly desirable.   But this advice should not be thought of as
definitive, finally authoritative or legitimating. As Hastings pointed
out[12] the nature and scope of an evaluation can all too easily be
limited to the capacity of the particular evaluation methods, techniques
or instruments with which an evaluation specialist is familiar.
Program participants should consider the extent to which the advice of
specialists and the evaluation processes and findings they propose will
contribute to the critical debate about the program and "program decision
making" (that is, whether the information and arguments collected will
help in guiding and refining action in the program).

In short, evaluation consultancy can help considerably in the planning
and execution of a self-critical evaluation. But program participants
should consider who is helped and how and when they will be helped by
particular evaluation methods and techniques. In the end, the community
of participants bear responsibility for the program, so evaluation plans
must be judged by reference to their impact on the community of participants
and on those who interact with it.

(b) Meta-evaluation

A program sponsor is a co-participant in the community of self-interests
formed by a program. But as an agency authority accountable for the
expenditure of its resources, it will generally need to be satisfied that
the evaluation arrangements proposed for a project or program are adequate
and appropriate. Program sponsors therefore have an interest in meta-
evaluation (the evaluation of evaluation), to determine whether a program
evaluation can meet the demands of the critical debate to which the
program is subject. This is in part an internal management question,
but it will naturally include an interactive element through which project

---

12   Hastings, J.T., loc.cit.

24

or program evaluators confer with program sponsors (who are often also sponsors of program evaluations). In exceptional cases, program evaluations may be formally evaluated by evaluation sponsors, but most often the meta-evaluation process will be in the nature of informal monitoring and interaction.

According to the principle of the self-critical community, the primary responsibility for program evaluation is "internal" to the program. Program or program evaluation sponsors may want to ensure, however, that adequate evaluation consultancy is available to the program community, and they may want to encourage some form of meta-evaluation which can monitor the responsiveness of evaluation efforts to the concerns of the program community and interested observers. Like the program evaluation function, the meta-evaluation function can be "devolved" from the sponsor to the program community in order to ensure that the range of concerns, perspectives and interests present in the program community are being considered in the critical debate about the program. It is possible that a meta-evaluation will guarantee that sponsors' interests in the evaluation are met at the expense of other interests in the same way that evaluation consultancy can serve some interests at the expense of others; thus, a meta-evaluation must be judged by reference to the same criterion as the primary program evaluation: how does it contribute to the improvement of the critical debate about the program as a whole, for the whole community of program participants and interested observers?

c) External evaluation

Following negotiation with other participants in the community of self-interests of a project or program, a program sponsor may sponsor distinct evaluation studies which are outside (or in addition to) the self-critical evaluation of the program. These evaluations will take two forms: external evaluation and independent evaluation.

External evaluation studies may be commissioned when the community of self-interests of a program wants advice, critical review or validation from substantive specialists, or when the judgment of recognised authorities in a field is necessary for a project, program or evaluation to be regarded as credible. In such cases, care should be taken in

negotiating an evaluation contract with potential external evaluators to see that the evaluation study respects the interests of the program community as a whole (as expressed in these principles). Most program evaluations undertaken today are of this form. Regrettably, they tend not to take account of the values embodied in these principles (e.g. rationality as reasonableness, autonomy and responsibility, community of self-interests, etc.). In consequence, they may further some interests in a program at the expense of others, impose a "foreign" perspective on the work of the program, or deny the authentic knowledge of participants. We should not be too xenophobic about this state of affairs, though: often such perspectives prefigure the views of the wider community outside the program (indeed, they often shape outside views). A healthy self-critical community should incorporate these external perspectives and where necessary correct them by reference to the concerns and circumstances of the program as a community with particular goals, terms of reference and contexts.

Just as it is a mistake to assume that an external evaluation represents the "true" perspective on a program (though it may aspire to objectivity, its very purpose will align it with particular interests in or around the program at the expense of other interests), it is a mistake to think of a self-critical evaluation as a complete amalgam or synthesis of relevant perspectives. Both kinds of evaluation are fluid and interactive, not susceptible of completeness or ruling definitively on the worth of a program. A self-critical evaluation aspires to awareness of the diversity of values and interests in and around a program and more conscious negotiated control of program development and evolution; an external evaluation aspires to awareness of particular values and interests, and to influencing program development or evolution in the light of these particular values and interests. The mistake is to think that either represents a unified or complete perspective on the program which provides an unequivocal basis for program action.

(d)   Independent evaluation

Sometimes a project or program will be so large or diffuse that its sense of being a community of self-interests or a self-critical community is sharply attenuated. It may be able to develop only a very poor sense

of itself as a whole. In other cases, program participants may prefer
a "specialisation of function" in evaluation, so that one person or a
small group take responsibility for the conduct of the evaluation
(though program participants will naturally continue to participate
in the critical debate). In such cases, an independent evaluator may
be appointed or commissioned in order to harness and refine the critical
process on behalf of the program. Such an evaluator may prefer to be
regarded as an evaluation "facilitator".

Moreover, when the community of a program embraces a wide variety of
viewpoints which must be articulated and explored before a joint
perspective can be reached about the value and meaning of the work, an
independent evaluation will be appropriate. Similarly, an independent
evaluation will be appropriate when the work of a program is particularly
controversial and a variety of value perspectives within or outside
a program must be considered before an evaluation can be regarded as
relevant and credible by audiences with differing value commitments.

Independent evaluations will often require the services of evaluation
specialists capable of dealing with conflicting value perspectives,
political pressure, complex theoretical conceptualisations, and the
real and difficult issues of curriculum associated with a particular
program. They may need to adopt refined evaluation procedures capable
of generating and maintaining negotiation among the conflicting
theoretical, practical, and organisational interests of those in and
around the program. It is necessary to take great care in negotiating
evaluation contracts with potential independent evaluators in order to
ensure that the contract (as much as the evaluation) respects the
values and interests of the range of participants in the community of
the program.

The principle of the self-critical community is a recognition of the
natural existence of self-reflection within a program, on the one hand,
and the natural critical debate around it, on the other. Such a principle,
may encourage those involved in project and program evaluation to be
"responsive" in the sense that Stake[13] uses the term. He says:

---

13   Stake, R.E. "To Evaluate an Arts Program", in R.E.Stake (ed.) Evaluating
      the Arts in Education: A Responsive Approach. Columbus, Ohio: Charles E.
      Merrill, 1975.

> *An educational evaluation is <u>responsive evaluation</u> if it*
> *orients more directly to program activities than to program*
> *intents; responds to audience requirements for information;*
> *and if the different value-perspectives present are*
> *referred to in reporting the success and failure of the*
> *program. (p.14)*

But in addition to this, such a principle may encourage evaluators to
see their work as part of a naturally-occurring process of evaluative
activity in a program, not distinct from it.

It would be possible to adopt alternative principles to this one.
On the one hand, curriculum programs could be evaluated solely by teams
of expert external evaluators, thus putting the validation function of
evaluation before all others. Or program sponsors could adopt a form of
evaluative activity based on their own perspectives of what projects and
programs should be, thus establishing the primacy of their own value-
frameworks (as seals of approval) in every program evaluation. But
neither of these principles will suffice. The cooperative nature of
curriculum development and the diffuse control of educational organisations
(with different participants having different sources of legitimacy --
teachers' professionalism, schools' autonomy, Ministerial responsibility
for State systems, parents and community roles in school councils,
students' rights, etc.) mean that curriculum evaluations must encompass
wider views than those of substantive experts or program sponsors'
particular predilections. Program evaluators simply cannot afford to
ignore the wider debate about a program in its social and educational
context.

Current trends in the history of evaluation have been significantly
influenced by the demands of project evaluation where outside groups of
evaluators have been called in to observe and evaluate curriculum
development work in order to provide external validation of the quality of
development. As a consequence, much recent evaluation literature reflects
an expectation that evaluations will be "objective", disinterested,
expert and validatory. But external evaluation cannot provide unilateral
validation. There is an older trend in evaluation based on school
accreditation, inspection and appraisal which is more organically related
to school curriculum work. But the techniques these purposes generated
are not well suited to the evaluation of innovative curriculum projects

or programs. The older tradition stabilises itself around the organisation of a school rather than around the organisation of a new curricular activity or product.

Project or program evaluation must be able to negotiate between the demands for curriculum validation and the conditions of schooling in different systems. Program evaluation cannot treat curricula as discrete products, to be considered as if they existed independently of their contexts of application, nor can it focus all its attention on the conditions in schools adopting particular innovations. The principle of the self-critical community recognises that innovations enter adopting systems by a process of negotiation; evaluation should facilitate negotiation by refining the critical debate.

The people who work on, use and sponsor a particular program for a natural focus for its evaluation activities; their work provides a natural forum for critical thinking about it. The principle of the self-critical community may encourage those associated with innovative programs to regard their natural evaluative work as a primary, not a secondary, evaluation function; accordingly, it is proper to expect that "internal", self-critical evaluations will provide the primary basis for judgments about the nature and worth of programs. Evaluation should not be regarded as a specialist activity tagged on to development to monitor and observe from a position of privilege (the outside observer) as if the interests which guided evaluation work were unrelated to the interests of those which guide the developers (that is, that there are no confluences or conflicts among their values and interests). Evaluation is interactive and reactive; it should not be construed as "objective" and outside the whole system of social relationships which constitute curriculum development programs in practice.

29

> *Evaluation processes inevitably affect the political
> economy of information in a program (the production
> and distribution of information about it). Because
> information and arguments justify or legitimise
> decisions, evaluation affects the distribution of
> power and resources in program situations. Program
> participants and interested observers live with the
> consequences of the use and abuse of evaluation
> information. An evaluation should have explicit
> principles of procedure which govern its conduct and
> its processes of information production and distribution.*

As suggested in the introduction to these principles, evaluation is often defined as "delineating, obtaining and providing information for making decisions and judgments about educational programs". Indeed, formal evaluation efforts may well be included among the management and decision making processes of a project or program. Though it has been an explicit purpose of these principles to widen that definition, it would be naive to assert that evaluation was not normally regarded as an important management and decision making tool. Evaluation processes thus link the generation of information and arguments about a program with the power to decide: those responsible for deciding the shape and conduct of a program, whether it should be implemented, or even whether it should be continued or discontinued will look to evaluation studies as sources of information and arguments when they make their decisions. Evaluation is thus inevitably a political process, affecting the flows of information in a situation and having life-consequences for those who inhabit it.

The point was made dramatically by ethnographer Harry Wolcott in a throwaway line at the Annual meeting of the American Educational Research Association in 1976. Discussing several papers on ethics and methodology in fieldwork, he remarked: "Some people define evaluation as the collection of data to guide decisions to continue, revise or terminate programs. If you were an ethnographer, how would you like your material to be used to continue, revise or terminate another culture?"

The production and distribution of information about people, projects and programs through evaluation must be regulated according to a principle of propriety capable of taking into account the moral, social and political consequences of information use and abuse. The evaluator must find

procedures appropriate to each context by which he or she can negotiate
the disputed territory between the public's "right to know", management's
"right to relevant information", and the individual's "right to discretion".
Even in cases where innovators are anxious to have their work more widely
known, or where teachers regard their work as exemplary, there may be
consequences of the release of information which may jeopardise their
future opportunities. Evaluators must treat seriously the problems raised
by the political economy of information production and distribution -- the
role of evaluation in the distribution of power in particular settings and
in the support or denial of already-existing power structures. It is not
sufficient to take a moralistic stance on open information, on privacy or on
the rights of sponsors: the production and distribution of information
inevitably affects the politics of the program situation and it is up to
the evaluator to find procedures which are defensible within the particular
context and technically feasible given the constraints of time and resources.

The principle of propriety in the production and distribution of
information implies that evaluators must set out their intended procedures
for information control in the form of an evaluation contract or a statement
of procedural aspirations so that everyone who becomes involved in the
evaluation process knows how the information is to be produced and distributed,
what risks are involved in cooperating with the evaluator, and what safeguards
exist against the misuse of evaluative information. Such procedures should
specify how information is to be collected, analysed, interpreted and
reported. It should indicate the status of the evaluator's interpretations
vis-a-vis the interpretations of program participants (including program
sponsors). The contract should make clear who will come to know what about
whom as a result of the evaluative process and its products. (The process
is just as important as the product in shaping the views of participants in
the evaluation). It should make it clear what procedures will govern access
to "data-sources" (people, records, events), the conduct of the evaluation
and the determination of its boundaries, the ownership of the evaluation
data and findings, the release of information, rights to publication,
confidentiality rules and mechanisms for accountability of the evaluation.
It may also be possible to specify safeguards against abuse of the intended
procedures (like rights of appeal or the sanction of denying the evaluator
further access to the situation).

Alternative principles can hardly be framed in terms of "impropriety": no one could accept the notion that an evaluation should use information improperly. The principle of propriety presented here does specify the rights of participants in a program to know how the information is to be used and controlled. It attempts to set up a model of equitable distribution of information based on the rights and obligations of all those involved in an evaluation study. Evaluations would operate in a spirit contrary to the present principle if they were an exclusive information service for evaluation sponsors rather than a service to a range of audiences associated with the program, if they used secret reporting, if they failed to take into account the diverse perspectives and interpretations of participants and evaluation audiences, or if they publish reports in forms suitable only for research audiences. The principle thus establishes a view of evaluation opposed both to the view that evaluation is an arm of the educational research industry serving some general ideal of truth or "the public interest", and to the view that it is a tool to be used in the service of bureaucratic responsibility[14]

Furthermore, the principle of propriety in the production and distribution of information establishes the view that evaluators have the responsibility to be aware of the consequences of information production and distribution and to respond in defensible ways by developing appropriate procedures for information control.

7   THE PRINCIPLE OF APPROPRIATENESS

> *Evaluation design is a practical matter. An evaluation*
> *must be appropriate to the program setting, responsive*
> *to program issues, and relevant to the program community*
> *and interested observers. An evaluation design must be*
> *renegotiated as the study progresses in the light of*
> *changing circumstances, issues and interests, and in*
> *the light of its own consequences (as they become apparent).*

The contemporary scene in evaluation theory and research abounds with evaluation models and approaches with a bewildering variety of foci and

---

14   See MacDonald, B. "Evaluation and the Control of Education" in D.A. Tawney (ed.) Curriculum Evaluation Today: Trends and Implications. London: Schools Council Research Studies, MacMillan, 1976; MacDonald distinguishes "autocratic", "bureaucratic" and "democratic" forms of evaluation in his political classification.

employing a diversity of specific techniques. While this variety and diversity must be acknowledged, evaluators and evaluation sponsors should not adopt an unconstrained eclecticism with respect to evaluation just because no dominant orthodoxy has emerged in the field. These principles and the value commitments they embody identify some forms of evaluation as unacceptable. To be acceptable, particular evaluations should embody the six principles previously presented, but they must also be appropriate to their objects. That is to say, evaluation studies must suit the curriculum projects, programs, processes or products to be evaluated and the contexts in which they appear. The design of an evaluation is a practical matter, depending on considerations of purposes; audiences; substantive issues raised by program theory, aspirations, organisation and practice; resources; issues of information control in the particular political economy of the program and its evaluation; relevant evidence; methods for data-collection; issues and approaches to analysis and interpretation; and modes of reporting.

Evaluators and participants in curriculum projects and programs must take all of these topics into account in designing or commissioning evaluation studies. The appropriateness of evaluation designs is a practical matter, not a technical or theoretical one. Decisions about the form an evaluation should take cannot be made by reference to the "internal logics" of evaluation models and approaches alone; such decisions must take into account the needs, preferences, obligations, circumstances and opportunites of those who will be most closely involved in the evaluation process (as evaluators, program participants, sponsors, evaluation audiences).

As in the case of the last principle, it is hardly possible to propose an alternative principle of "inappropriateness". But inappropriate evaluation designs are often proposed for the evaluation of curriculum projects and programs. Such designs are ones which suffer from "methodological tunnel vision", employing evaluation models dogmatically or inflexibly when more sensitive attention to the critical debate about a program or the circumstances of its operation would suggest a different approach. Evaluation designs are also inappropriate when they fail to serve those most closely involved in the work of a program, reporting instead only to sponsors or research audiences. These audiences have a legitimate claim for evaluative information, to be sure, but evaluations frequently fail to serve the needs and interests of those most directed affected by the work.

If the evaluators of curriculum projects and programs take seriously the thrust of the definition of evaluation proposed at the beginning of this paper -- that it is the process of marshalling information and arguments which enable interested individuals and groups to participate in the critical debate about a specific program -- then it is less likely that they will err on the side of inappropriateness. Appropriate evaluations will take into account the social and contextual conditions under which educational programs operate and include a meta-evaluation component: the evaluation will thus include an element of self-reflection which allows those involved with the evaluation and the program to monitor its effects on program development and evolution and on the social life of the program as a community. The aim of this self-reflection is to treat the appropriateness of the evaluation as problematic and dynamic, not as something which can be decided once and for all at the design stage. It is to recognise that evaluation programs, like curriculum programs, are negotiated between interested individuals in the light of contemporary circumstances and opportunities and renegotiated in the light of their consequences.