ABSTRACT
                Evaluation as a political arbitration entity is
discussed in the case of a multilateral literacy development project
in the fourth year of operation in a Third World country. An external
evaluation team was invited to evaluate the project when conflict
appeared between the funding agency (A) and the technical agency (B)
over a project-related issue: the extension of an expired
Instructional Materials Expert (IME) position. Agency A felt that IME
tasks were completed, while B felt that the post extension was B's
decision. The recommendations of the evaluation team were
organizational and dealt with personnel issues: (1) an exit date to
transfer project responsibility from A and B to government agents;
(2) a change from project supervision to government advisor for a
technical agent in preparation for B team exit; and (3) a four-month
government funded extension of the IME position. The paper concludes
that evaluation is political in its genesis, political as a process,
and political in its effects. This means that evaluators must begin
with collecting political data of relevance to the evaluation action,
and must learn to take account of political intelligence in the
process of generating, interpreting, and utilizing technical
information. (PN)

EVALUATION AS ARBITRATION:

EXTERNAL EVALUATION OF A MULTILATERAL DEVELOPMENT PROJECT

IN A THIRD WORLD COUNTRY

By

H.S. Bhola
School of Education
Indiana University
Bloomington, Indiana, 47405

October 17, 1985

1

EVALUATION AS ARBITRATION:

EXTERNAL EVALUATION OF A MULTILATERAL DEVELOPMENT PROJECT

IN A THIRD WORLD COUNTRY

By H.S. Bhola

Evaluation   technology confounded by politics. As a
technology, evaluaticn seeks to generate needed information about
contexts, inputs, processes and outputs of development programs
and projects, systematically, objectively and with exactitude,
for such information to be used to improve both the efficiency
and the effectiveness of programs and projects being evaluated (1).
But as Cohen points out, evaluation is political as well:
"Evaluation is a technique for mersuring the satisfaction of
public priorities; to evaluate a social action program is to
establish an informtion system in which the main questions
involve the allocation of power, status, and other public goods"
(2).

It is now quite well understood that there is more to
evaluation than doing science. The evaluation process iJ indeed
entangled with individual politics, organizational politics, and
the politics of the macro systems within which it is conducted.
As Carol Weiss has reminded us, programs being evaluated are,
after all, creatures of political decisions and evaluation
reports do enter the political arena. Evaluation itself has a
clear political stance (3). The decision to commission an
evaluation of a program or a project is a political decision; and
the evaluator, in accepting to conduct the evaluation, makes an

2

implicit commitment to the existing social order. The evaluator also accepts the intervention model embedded in the program structure; and then makes a political choice as the experimental or the participative methodology of evaluation is selected. There is a clear relationship between the power structure and the choice of indicators of change; and those innocent looking tests are ultimately connected with the social mobility system. Politics gets involved in the choice of sites for the evaluation exercise; and in the choice of criteria and standards to judge the success or the failure of the program. The data from the field come wrapped in values; and thus the so-called hard facts are soft at the core. Again, politics gets involved, first, in the interpretation and, then, in the utilization of results. And, finally, there is the politics of "the evaluation of the evaluation" itself (4).

Unfortunately, evaluation theory has yet to catch up with the realities of evaluation in the real world. In a recent review of evaluation theory, Stufflebeam pointed out that early evaluation models all "tended to equate evaluation with the use of some preferred techniques" (5). The "participatory evaluation" model, more than the "advocacy evaluation" model, can respond to political realities. Yet, Guba and Lincoln's remark that sociopolitical context has been excluded from evaluation "as beyond the pale of 'objectivity'" remains largely true (6). Cronbach's thesis that "a theory of evaluation must be as much a theory of political interaction as it is a theory of how to determine facts" (7) throws a conceptual challenge yet to be taken up by evaluators in right earnest.

3

This paper does not pretend to pick up the theoretical challenge and accomplish the theoretical task in one fell swoop. But we do hope to take a step towards a political theory of evaluation. This theoretical step is taken on the basis of the evaluation, recently undertaken by the author, of a multilateral development project in the fourth year of its operation and located in a least developed Third World country (8).

The Case:
Evaluation as Arbitration

The case presented below is a special case of the politics of evaluation. It demonstrates, I think, a special use of evaluation -- evaluation for purposes of arbitration of a dispute between two multilateral institutions of technical assistance jointly responsible for the same development project.

Two multilateral technical assistance agencies, UN(A) and UN(B), were jointly responsible for a "literacy for development" project (Litdev) in a Third World country (Xland). UN(A) provided funding for the project and, through its resident representative in Xland, supervised and monitored the project. UN(B) as a technical agency was responsible for the professional implementation of the project. As such, it selected and appointed all professional staff; and, throught them, was responsible for all professional and technical decisions as well as for resource allocations within the agreed upon budgetary limits. The UN(B) did not have a resident representative of its own in the country and was indeed represented to the Government of Xland through the resident representative of UN(A).

4

The arrangement worked quite well, until conflict appeared between the two agencies over some project-related issues. It is at this time that the agencies involved invited an "external" evaluation team to evaluate the project. The team can justly be called external because none of the four members of the evaluation team was associated with the project as a functionary, nor was any one of them currently employed by UN(A), UN(B) or the Government of Xland.

The evaluation team would be a tripartite team: each of the three institutions involved, UN(A), UN(B) and the Government of Xland would appoint a member. The fourth member of the team would be an outside literacy specialist with experience in evaluation in the Third World and would be appointed by UN(B). The team would be led by the member appointed by UN(A). The team would stay in the country for two weeks and follow the procedures and methodologies established by mutual agreement bewteen UN(A) and UN(B) and accepted by the Government of Xland.

The terms of reference and the themes of concern

To anticipate the story somewhat, it was intriguing to note that there were two streams of consciousness intermixed in the communication behaviors of everyone involved in the evaluation project. At one level, they talked about the need for evaluating the project. At another level, there was the search for arbitration of a dispute.

This ambivalence was quite clear in the terms of reference document provided to the evaluation team. While the document went to some length in describing the professional needs of the

project that the evaluation team had to meet, the document did betray the theme of its real concern. For instance, the terms of reference document began by describing the objectives of the evaluation to be to determine how effectively the intended purposes of the project were being attained and how effective the project was likely to be in the future to contribute to both sectoral and national purposes of Xland. The evaluation was also to identify factors that had facilitated or deterred the achievement of intended objectives in the past; and how those factors might be manipulated in the future for ever more effective actions. This was the language that provided the technical rationality for the conduct of the evaluation exercise.

Hidden under these pretensions of technicality however, there was a clear theme of a conflict in search of arbitration. The Litdev was a two-man project as far as the foreign technical assistance personnel were concerned. One of the posts had expired a few months earlier and was now being extended month by month from extra-budgetary sources. UN(B) as the technical agency responsible for the project wished to extend the post because, according to them, some important tasks still remained to be accomplished. However, UN(A) was not convinced. The various agencies of the government of Xland came to join the issue on different sides, and two opposing coalitions of organizations were formed.

As UN(A) saw it, the tasks of the Instructional Materials Expert (IME) were all done. The IME had been on the Litdev project for three years, during which time, curricular decision

6

had all been made, materials for use in the Litdev had been designed and produced. No doubt, there were some additional instructional tasks that remained to be achieved , but those had been anticipated, at the project planning stage, to be done by short-term consultants with specialized technical skills. The implication was, of course, that the present IME did not have those specializations at the level required. The UN(A) had consulted with the team leader of Litdev, a UN(B) appointee, who had agreed with their reading of the situation. The local authorities took too long to come on the side of the IME and in making the statement that his post did need continuation. In any case, the money for the continuation of the post had run out and UN(A) disliked being put in a position where they were forced to pay the IME from funds meant for short-term consultants. They resented all the cable traffic they had to receive and to generate. As far as they were concerned, it was not right to personalize an official issue. Some of the agencies of the Government of Xland were also of the same view. Finally, the personal egoes of many of the UN(A) staff were bruised when they were characterized by UN(B) as administrators who should not be interfering in professional matters.

UN(B), on the other hand, felt that as the technical agency responsible for the implementation of the project, it was for them to decide whether the post of the IME needed to be extended or not. The UN(B) HQ took their local team leader to task for having played in the hands of the local office of the UN(A) and for having suggested that the tasks of the current IME had indeed been completed; a local conterpart had already been trained; and

7

that the need for short-term specialist consultants was paramount to the continuation of the IME. UN(B) had found money from a third party, an international foundation dealing with development, but the gesture had backfired. The local treasury department of Xland took exception to UN(B) commiting the Government of Xland to contractual arrangements with any outside institution or government without prior consultation with and clearance from the government.

In the ten months prior to the external evaluation here reported, two internal tripartite reviews by UN(A), UN(B), and the Government of Xland had been conducted but no agreement had emerged. The arbitration role of the present evaluation was quite clear to everyone, thought no one talked about it directly and forthrightly.

The two faces of evaluation

While the whole of the project team was aware of the politics of the evaluation, this author had a special theoretical interest in the evaluation (See No. 8 of Notes and References below); and did in fact develop, as systematically as possible, two profiles of the project: (i) a technical profile and (ii) a political profile.

It is not within the scope of this paper to include the details of either the technical or the political profile of the project. What we need to point out is that a political profile is developed by asking political questions in regard to the existing politics at the individual and organizational levels and at the level of the social context. As pointed out earlier, in this

8

9

particular evaluation, there were no formal instruments that were
administered. The data were all collected through observation,
interviews and discussions in individual and group settings. It
was possible for the author to ask political questions especially
in individual interviews. The fact that this evaluator had
worked in the country before in other educational and development
oriented projects helped in establishing rapport and trust.
Perhaps, this is the best way to go: technical questions may be
included in formal questionnaires and interview schedules, while
data for the political profiles are collected more informally.
The point to be made here is that political data must be
collected in ernest and not dismissed as so much back-biting and
mud-slinging deserving of no serious attention.

Formulation and presentation of recommendations:
Concerns for arbitration met

The concerns for arbitration were obvious in the way
recommendations were developed and presented to the various
stakeholders. At the stage of data collection all parties to the
dispute had been interviewed and all the various levels of the
enterprise had been covered from the village to the central office.

The recommendations of the evaluation team, quite
expectedly, were mostly organizational and dealt with personnel
issues. It had indeed been made quite clear in personal
briefings, by both UN(A) and UN(B), that they wanted to have a
clear resolution of the IME issue and wanted the recommendations
of the evaluation team to be clear, unequivocal and such that
could be acted upon both professionally and administratively.

9

10

Technical recommendations made by the evaluation team were
sensible but quite general. The leadership of the country was
asked to make a clear political commitment to the promotion of
literacy at the national level. The need for strengthening the
system of advice through the activation of a national board for
literacy was recommended as also the need to strengthen and
support the system of administration and coordination from the
center down to the village level. The team's report recommended
that due attention be given to the establishment of linkages
between the literacy agency and other development and extension
agencies of the government; and, finally, much greater emphasis
on the training of personnel, at all the different levels of
responsibility, in a variety of training settings such as courses,
workshops, conferences and on-the-job training, was recommended.
All these recommendations were offerings on the altar of technical
rationality.

The recommendations on the administrative aspects of the
program were much more important and concrete and served the
arbitration needs of the project. The first recommendation
related to the fact that the project which had been in operation
for almost four years should have an exit date by which time
there should be a transfer of responsibity from UN(A) and UN(B)
to the agents of the Government of Xland (9). This exit date was
set for December 31, 1985, sixteen months away from the date of the
external evaluation.

The chief of party, one of the two-man UN(B) project team,
was asked to assume a new role as planning, programming and training

10

adviser instead of the current role which involved considerable amount of administrative duties including watchdog functions in regard to the properties contributed by UN(A) and UN(B) and superivion of the work of IME. In other words, the role of the chief of party was changed from supervisory to advisory to the government. The main purpose of this recommendation was, of course, to take responsibility away from the UN(B)'s official and give it to the government; and prepare the UN(B) project team for exit. It also freed the IME from the authority of the chief of party at least during the last few months of his stay at the project. The post of the chief of party was authorised until the end of the Litdev project. His particular role on the project and his demonstrated competence on the job stood him in good stead.

The post of the present IME was extended for four months only, exclusive of accrued leave. The government was requested to explore extra-budgetrary resources to fund the extension of the IME's post so that funds budgeted for consultant services are not used, if at all possible. In other words, the government was requested to let UN(B) off the hook for having requested and obtained funding in behalf of the government from a third party, without prior consultation and without their formal request for such assistance.

It should be clear from the nature and content of the recommendations made that the evaluation did indeed have an arbitration function. The way the recommendations were presented first in the "draft" form to a meeting of all stakeholders; and the counselling provided by the team leader to the IME who may

11

have felt to have lost the battle, were all in the nature of
negotiating and healing.

Lessons from the case

The case presented above does lend support to the
assertion made in the beginning of the paper that evaluation is a
technology confounded by politics. Evaluat ons are indeed part
of the prevailing politics as various stakeholders seek power,
authority, advantage or accommodation within their social
settings.

While social evaluation is always political, it is by no
means always for arbitration and negotiation. On the contrary,
evaluation often can (and does) generate disagreement and feed
conflict, as "facts" of the matter brought out or missed by
an evaluation are used to support arguments and counterarguments.
Thus, the use of evaluation for arbitration is just one of the
many aspects of the politics of evaluation. We chose to bring
this case to the present audience because we found the
realization of evaluation as arbitration personally intriguing
and theoretically interesting. The question now is: Was this a
unique case, a universe of one, unlikely to be repeated
elsewhere, or should we anticipate multiple instances of the use
of eval ation as arbitration in the world of education and
development?

In our view, the use of evaluation as arbitration is and
will continue to be frequent. In a paper that I published many
years ago, at the end of a two-year tour of duty with an UN-
affiliated agency, I had suggested that the delivery of education

within the context of bilateral and multilateral technical
assistance was indeed an act of "educational diplomacy" (10).
The diplomatic frame establishes the limitations as well as the
possibilities of educational technical assistance. The
limitations are real but do not render professional educational
work impossible within the context of technical assistance.
Indeed, an understanding of the limitations may make the
actualization of possibilities more effective.

The same seems to be true about educational and social
evaluation. The conduct of evaluation within the context of
bilateral or multilateral technical assistance is also an act of
"evaluational diplomacy." The diplomatic frame establishes the
limitations and possibilities of evaluation. Again, the
limitations do not abhort the possibility of professsional
evaluation. Indeed, the awareness of the political context may
make the evaluation more useful than when no attention is paid to
poltical realities and evaluators pretent to stand outside
politics. Those of us who have worked with evaluations of
educatic .¹ and development programs in the technical assistance
context in the Third World know, of course, that evaluations do
evaluate processes and effects. Projects are reconceptualized or
reorganized in the field, and some of them are terminated
prematurely as a result of evaluations.

In making the preceding statements, we are, of course,
making epistemological claims in regard to the nature of facts in
social life and the methods of generating knowledge for use in
the improvement of the human condition (11). We are assuming, on

the basis of fast accumulating experience, that in responding to the political realities of evaluation, the evaluator need not purposely engage in selective perception, or deliberately forget how to count, or ask trivial questions while real issues are avoided. Data need not be fabricated, nor is is necessary to offer misleading interpretations (12). It is indeed possible to be both "scientific" and "moral" within the political context of our work. Evaluators can learn first to see and then to speak truth to power (13).

Toward a political theory of evaluation

Current methodological debate in research and evaluation is coming to the view that a good theory of evaluation has to be a logico-dialectical theory. It must include the logic of systematic generation of information that is credible; and, at the same time, must be built upon the realization that both the processes of information generation and information utilization involve a complex of dialectical relationships between the evaluator, the evaluand, and all other stakeholders that have a stake in the evaluation and in the life of the program being evaluated. A useful theory of evaluation must enable the evaluator to "frame the flux" of the real world of social action. It must be able to accommodate both norms and counternorms, and both principled actions and selfish private needs.

While the creative ideology of the new theory of evaluation must be logico-dialectical, the essential nature of the theory has to be political. Evaluation is political in its genesis, political as a process, and political in its effects.

14

15

Clearly, this means that evaluators must always begin with
collecting political data of relevance to the evaluation action;
and must learn to take account of political intelligence in the
process of generating, interpreting and utilizing technical
information.

It is a challenge to the best of the evaluators in the
business today to develop orientations, steps, and procedures
that will ensure a symbiosis between the technology and politics
of evaluation, ensuring that evaluation is of and above
politics at the same time and that it contributes effectively to
individual and social praxis.

NOTES AND REFERENCES

1.   H.S. Bhola, Evaluating Development Training Programs.
Bonn, FRG: German Foundation for International Development (DSE),
1982.   [ERIC Document No. ED 238 651.]   See also H.S. Bhola,
Evaluation Planning, Evaluation Management and Utilization of
Evaluation Results within Adult Literacy Campaigns, Programs and
Projects. Bonn, FRG: German Foundation for International
Development (DSE), 1982 which draws due attention to the politics
of evaluation. [ERIC Document No. ED 221 759.] An earlier
publication addressed to literacy workers is H.S. Bhola,
Evaluating Functional Literacy. Amersham, U.K.: Hulton
Educational Publications (Literacy in Development — a series of
training monographs, edited by H.S. Bhola, and commissioned by
the Unesco/Iranian International Institute for Adult Literacy
Methods), 1979.

15

2.   D.K. Cohen, "Politics and Research: Evaluation of Social
Action Programs in Education," Review of Educational Research,
1970, 40:213-238.

3.   Carol H. Weiss, "Evaluation Research in the Political
Context," in Elmer L. Struening and Marilynn B. Brewer (eds.),
Handbook of Evaluation Research (The University Edition).
Beverly Hills, Ca.: Sage Publications, 1983.

4.   Some of the political aspects of evaluation listed in
the paper come from Gigeon Sjoberg, "Politics, Ethics and Evaluation
Research" in the Handbook of Evaluation Research quoted in
Note 3 above. Others came from the author's personal suffering
from the politics of evaluation.

5.   Daniel L. Stufflebeam, "A Review of Progress in
Educational Evaluation ," Evaluation News, Vol. 3, No. 2, May
1982, pp. 15-27.

6.   Egon G. Guba and Yvonna S. Lincoln, Effective Evaluation.
San Francisco, Ca.: Jossey-Bass Publishers, 1981.

7.   Lee J. Cronbach and others, Toward Reform of Program
Evaluation. San Francisco, Ca.: Jossey-Bass Publications, 1980.

8.   The author was part of a four-member evaluation team that
visited the Third World country in question. There was, of
course, the practical obligation to accomplish the evaluation
task as part of the evaluation team. But the author, as an
academic, had a keen theoretical interest in the politics of
evaluation. He, therefore, used the opportunity provided by the
visit to do an evaluation over and above the contracted
evaluation, to test some of the ideas developed in his earlier

17

work in evaluation of development programs in the Third World.
He thus conducted two simultaneous evaluations: one for the team
and one for the self. The present paper is a product both of the
team evaluation and the author's "private" evaluation of the
project in question.

9.   A comprehensive review of a number of UN projects in Third
World countries had come up with the need for projects to have an
exit date and to prepare systematically for the transfer of
responsibility to the local officials. See UNDP/UNESCO,
Educational Innovation and Reform (Evaluation Study No. 7),
New York and Paris, no date. This publication is loaded with
useful insights for the effective delivery of technical
assistance in the Third World.

10.   H.S. Bhola, "Limitations and Possibilities of Educational
Diplomacy: A Theoretical Framework," Viewpoints (A special issue
on International Dimension of American Higher Education), Vol.
47, No. 5, September 1971.

11.   Donald Polkinghorne, Methodology for the Human Sciences:
Systems of Inquiry. Albany, N.Y.: State University of New York
Press, 1983.

12.   Another aspect of the politics of evaluation is alluded to
in H.S. Bhola, "Building a Built-in Evaluation System: A Case in
Point." Paper presented to the 1984 Joint Meeting of Evaluation
Research Society and Evaluation Network held in San Francisco,
Ca., October 11-13, 1984. [Accepted in ERIC. Document No to be
assigned.]

13.   Aaron Wildavsky, Speaking Truth to Power: The Art and
Craft of Policy Analysis. Boston: Little, Brown and Co., 1979.

17

18