The Evaluator's Role

Lyle F. Bachman
University of Illinois at Urbana-Champaign

One of the most common problems in program evaluations is the friction that may arise between the evaluator and practitioners--program developers and teachers. Sources of friction include differences in personalities, interests, and frames of reference of evaluators and practitioners, the characteristics of the program or institution, and differences in perceptions regarding the procedures and uses of the evaluation. The program evaluator's most important responsibility is to minimize the opportunity for such conflicts to arise. This can be best accomplished if the evaluator assumes the general role of consultant, or facilitator of the program development process, and follows several specific courses of action to assure good rapport and cooperativeness with practitioners.

I'd like to talk with you about friction. Not the friction that you get when you rub two sticks together, but the kind of friction that too frequently occurs when practitioners and evaluators come into contact. I've experienced this sort of friction in many of the programs with which I've been involved, and I'm sure that most of you have experienced the same sort of friction, either as a practitioner in a program or as an evaluator of one. I'd like to share with you some ideas about the nature of this friction and about the kinds of roles and actions the evaluator can take to lessen it. Finally, I'll discuss some of the qualities and abilities of evaluators which form the basis for open, productive cooperation with practitioners, and for fair, useful evaluations.

But first, in order to illustrate some of the roles that can lead to friction, I'd like to introduce you to some people that we've all met at one time or another. Our first guest is Judge Learned Nose, of the blue-ribbon panel of expert evaluators:

> "It is the finding of our evaluation that your program has manifestly failed to provide the services for which it was developed. Specifically, students in your program have fallen below national norms on the **Test of Exaggerated and Fractured Language**, and have been required to spend long hours studying. Indeed, they have become very unhappy in the program. It is therefore the judgment of this panel that this program be discontinued immediately. Furthermore, you and all those responsible for its development will be required to cooperate fully with the new development team, which will have complete authority to throw out all your work and begin a new program."

Shades of the Inquisition! Next we have Inspector Seymour Faults, of the State Department of Evaluation:

> "Good morning. I'm Inspector Faults, Evaluation. [flashing badge]. I'd just like to ask a few questions about your program, so you don't need to feel so nervous. Do you mind if I take notes? Thanks. First, could I see your performance objectives? That's right, performance objectives. I see...[makes a note of this]

Have you included processes in your plan? What plan? [More menacing] Have you kept complete and accurate records of classroom interactions and student progress? Just the facts, sir, just the facts. [Exasperated now] How can you expect your subjective judgments to be reliable and valid? [Resigned to the inevitable] I can see that our people are going to have to come in and administer some criterion-referenced achievement tests. Can you set this up for tomorrow morning? And in the meantime, I'll need to see all your records by this afternoon. I guess that's all for now!"

Glad he's not on my case! Next we have Prof. Randolph M. Sample, of the Omniscient Ivory-tower for Studies in Evaluation:

[In a very condescending tone] "Well, from the little you've told me, the only acceptable design which will let you look at all these effects as well as the interactions among the main effects and your covariates would be a 4-way factorial, with counterbalanced orders of tests. This would absolutely require, of course, random selection and assignment to treatment groups. With ten covariates, you'll need at least 300 subjects per cell, or about 4,800 in all. The analyses should be pretty straight-forward. A multivariate analysis of variance and covariance, probably using a hierarchical step-down regression procedure should suffice."

Nothing like a clear, direct answer, is there? Finally, we have Mr. Gladly Teche, of Community School District 316:

[Tone of frustration] "Those eggheads have no idea what we're trying to do. We know that all we have to do is let the students communicate, but they expect us to have this all broken down into specific objectives--in advance, no less! They rearrange all our classes and schedules for their convenience, and then they expect us to let them come waltzing into class poking their noses into everything. One of them even wanted me to give up 15 minutes of precious class time to let him interview some of my students! And then they'll come back and give us a bunch of statistics to show that we did everything wrong. I wish they'd just let us get on with our teaching."

While these examples are a little exaggerated (I hope), they do serve to illustrate the kind of communication gaps that frequently lead to friction between evaluators and practitioners. They also point to a variety of sources of this friction. Perhaps the most realistic thing for an evaluator to keep in mind throughout an evaluation is this paraphrase of Murphy's law:

If friction can occur, it will.
No evaluation is ever as smooth as it seems.

One of the evaluator's most important responsibilities, therefore, is to eliminate, or at least minimize the opportunity for friction to arise. This can be done most effectively, I believe, if the evaluator assumes the general role of consultant, or facilitator in the development process. The evaluator will also assume a variety of specific roles during the evaluation. These will depend on how much the evaluator directly intervenes in the program, and the types of activities he or she undertakes. Does the evaluator assume the role of advocate, for example, or of sounding-board? Is he interested primarily in giving the practitioner feed-back after the evaluation is complete, or in involving the practitioner in the process of evaluation?

In order to successfully avoid friction, what abilities and qualities should the evaluator possess? Practically any list of professional capabilities of the "ideal" evaluator reads like a "com-

bination of the Boy Scouts' laws, requirements for admission into heaven, and the essential elements for securing tenure at an Ivy League college."[1] While very few people have all these qualities, most of us can either acquire or cultivate those we lack. Indeed, just being aware of what they are can have a positive influence in minimizing role conflicts and friction.

I'd now like to turn to the nature of the friction between practitioners and evaluators. In her book, *Evaluation Research*, Carol Weiss discusses several sources of friction, as well as several issues in evaluation that frequently lead to friction.[2] Weiss groups sources of friction into three types: 1) differences in personalities, interests and frames of reference, 2) differences in roles, and 3) characteristics of the program or institution.

First let's look at *differences in personalities*. Some people believe that practitioners and evaluators have different basic personality traits:

Practitioner	Evaluator
warm, outgoing, committed to action	cool, reserved, uncommitted

Where these personality differences exist, practitioners may find it difficult to relate to the evaluator. Practitioners and evaluators may also have different interests:

Practitioner	Evaluator
people, specifics, day-to-day activities	ideas and abstractions, generalizations, long-term acquisition of knowledge

Where these differences exist, practitioners are likely to see the evaluator's activities as irrelevant and of little immediate use.

A third area where practitioners' and evaluators' personalities may differ is in their frames of reference:

Practitioner	Evaluator
individuals, individual differences	groups, trends, similarities and differences across groups

If differences such as these are present, the practitioner may feel that the evaluation glosses over crucial differences. For example, it is of little consolation to a teacher to know that the average score of her class has increased as long as several individual students are still not succeeding.

Another potential source of friction lies in the *difference between the roles of practitioner and evaluator*:

Practitioner	Evaluator
advocate	skeptic
belief	questioning

To illustrate how this difference in roles can lead to conflict, let me share the following experience with you. As head of a team of persons working on a curriculum development project at a university in the Middle East, I initially assigned both development and evaluation activities to everyone. Thus, each member of the team was responsible for both developing and evaluating a given piece of the curriculum. Everyone got along fine under this arrangement. However, as the work progressed, the members of the team began to be concerned that their commitment to the program might be biasing their evaluation of it, and we therefore decided to assign one person to do all the evaluation. Although one of the most capable persons was the unanimous choice for this task, the relationship between this person and the rest of the team deteriorated almost immediately. He was seen as a turncoat because he now questioned everything he'd accepted and believed in before.

Finally, there may be *institutional characteristics* that can lead to friction. If there is a history of internal conflict, it will be difficult for the evaluator to remain neutral. And even if he does so, he is likely to be perceived by both sides as biased in favor of the other. If there is a high level of grievance and low satisfaction, the evaluator can easily become a surrogate channel for grievances, making it difficult to obtain objective information from practitioners. If areas of authority are ambiguous or fragmented, the evaluator is likely to be perceived as a spy for the person or persons who initiated the evaluation. In situations such as these, the potential for friction is high, and the evaluator must identify the sources of the friction as early in the evaluation as possible.

As if this were not enough, there are several *issues* which can lead to friction. Every evaluation is likely to involve the collection of data. This may require the administration of tests, interviews and other means of collecting information about students, as well as the updating of records to provide complete and accurate information about the program. Practitioners may find this extra work burdensome and may not see the relevance of this information to their needs. Or they may question the evaluator's procedures and the specific instruments used. It is therefore important for the evaluator and practitioners to discuss the rationale and procedures for the evaluation in advance. This not only informs the practitioners of the procedures to be followed, but it also helps to identify potential problems before they occur.

Another issue that can lead to friction is the selection of subjects and their assignment to groups. For evaluation purposes, specific grouping patterns may be important. For example, it may be of interest to compare homogeneous and heterogeneous grouping, or to look at particular groups according to their proficiency, native language, or academic discipline. Sometimes it is necessary to assign students randomly to treatment and control groups. Practitioners may feel that it is better to use a grouping pattern that they know will work effectively, and may see no reason to try others. Or they may feel that the use of a control group constitutes the denial of instruction that they feel is vital to their students' academic success. Again, discussing the evaluation in advance can help avert serious problems and alleviate these concerns.

A third issue which can lead to friction is regarding the use of feedback from the evaluation. The practitioners want to use results immediately for improving the program. In some cases, however, it may be necessary to keep the program stable over a period of time in order to study its effectiveness. As a result, the practitioner may see the evaluation as irrelevant and of little use.

From all of this, it may seem as if friction and conflict are inevitable in any evaluation, but this is not necessarily so. Indeed, in all successful evaluations, the rapport and cooperativeness established between the practitioner and the evaluator permits either the avoidance of conflict or its satisfactory resolution. The key to this lies in the perspective or role taken by the evaluator. "Good" evaluators, I believe, see themselves not as judges, detectives or statisticians, but rather as development consultants. Their role is to facilitate changes that will lead to the increased effectiveness of the program or institution.

As a consultant, the evaluator may assume a variety of specific roles, and pursue different actions. An example of an indirect role is that of objective observer, reflector, or sounding-board.[3] In this role, the evaluator asks reflective questions, and helps the practitioner confront and clarify the issues of the evaluation. The evaluator communicates none of his own ideas or beliefs, but lets the practitioner take responsibility for the decisions made and directions taken.

A slightly more direct role might be that of joint problem-solver.[4] In this capacity, the evaluator operates as a peer with the practitioner in planning and conducting the evaluation.

Together, they isolate and define factors which need to be considered in the evaluation, identify and weigh alternative courses of action, and identify resources which can be employed in the evaluation. In this role, the evaluator is able to help the practitioner better understand and conduct his own evaluation.

A very common role assumed by the evaluator is that of information specialist, or "expert."[5] I would like to point out that this role is not always the evaluator's choice, but that it is often forced upon him by the practitioner. In this role, the evaluator assumes direction of the evaluation, defining its objectives, procedures and utilization. There are two serious problems with this role. First, the likelihood of an inadequate evaluation may be increased because of the evaluator's limited perspective and inability to consider the full range of alternatives open to the practitioner. More serious, I believe, is the problem that this role leads to an increased dependence of the practitioner on outside "experts."

In the course of any given evaluation, the evaluator is likely to take on a wide range of roles. In each, his effectiveness will be determined by how well he can keep his perspective as a consultant, either assisting the practitioner in the process of the evaluation, or working with the practitioner as a peer. With this in mind, I'd like to refer to some specific courses of action suggested by Carol Weiss that the evaluator can follow to assure good rapport and cooperativeness with practitioners.[6]

Involve practitioners in the evaluation. Early planning with the practitioner has several payoffs: 1) practitioners gain a better understanding of what evaluation is all about--its purposes and procedures. This takes the mystique out of the evaluation, and removes much of the threat many people feel from evaluation; 2) practitioners can provide information and ideas that will make the evaluation more relevant to their needs and which can prevent the evaluator from making *faux pas* and unacceptable requests of the program staff; and 3) practitioners are more likely to cooperate if the evaluation is seen as their responsibility as well. In addition to planning, continued involvement of practitioners throughout the evaluation will facilitate communication and maintain the rapport established. Finally, involving practitioners in the interpretation of the results of the evaluation is essential, not only to assure the broadest and most reasonable interpretation, but also to assure that the results are not misinterpreted.

Minimize disruptions. Follow established procedures where possible and limit demands to indispensable issues.

Emphasize theory and program, not people.

Provide timely feedback of relevant information whenever possible.

Establish clear role definitions and authority structure. Define and assign various tasks and activities, and establish clear, reasonable lines of authority for resolving differences, *in advance of the evaluation.*

At this point you may be asking yourselves what qualities and abilities enable the evaluator to assume these various roles. (Or, perhaps, why anyone would be crazy enough to want to!) Lippitt and Lippitt discuss a broad range of qualities possessed by the "ideal" evaluator.[7] These qualities can be grouped into three categories: conceptual sophistication, a stable personality, and strong interpersonal skills. The conceptual sophistication required of an evaluator includes 1) the knowledge of evaluation methodology and skill in the use of research tools, including research design, measurement and statistics, as well as the perspective that these are tools and not ends in themselves; and 2) skill in diagnosing and analyzing areas of concern, identifying different alternatives, and implementing different approaches.

The stable personality necessary for conducting evaluations includes 1) maturity, with a self-awareness of areas of weakness and confidence in areas of strength; 2) a sense of professionalism, with integrity and responsiveness to the needs of the practitioner; 3) open-mindedness, with flexibility and tolerance for ambiguity; and 4) patience, with a high frustration level.

The interpersonal skills needed include 1) communication skills, with high ability for listening, observing and reporting; 2) the ability to work with groups; 3) the ability to relate effectively to a wide variety of people; and 4) a humanistic value system, recognizing the importance of individuals and that technology and efficiency are means, not ends, and placing trust in people and the domocratic process.

From this list, it may seem that the evaluator must possess superhuman qualities. This, however, cannot be so since, contrary to popular belief, evaluators are human beings. Indeed, the evaluator's most important qualities--humanistic values, communication skills and patience--are those which we all possess to some degree, and which contribute to whatever success we may achieve as practitioners. And the other skills can be acquired by anyone through diligence and experience.

In this paper I have attempted to face up to an all too common problem in program evaluation--friction between practitioners and evaluators. I have described some of the issues which can lead to friction and discussed some actions evaluators can take to avoid or resolve this friction. The key to this, I believe, lies in the evaluator's ability to take on some of the qualities and concerns of the practitioner and to work with him as a peer.

Evaluation is an exciting and challenging undertaking, as is program development. The role of the evaluator is demanding, as is that of the practitioner. The time for program development and for evaluation is always short, and the resources few. Can we, as practitioners and evaluators afford not to draw on those qualities we have in common--humanistic values, professionalism, patience--to utilize our differing skills, interests and perspectives in complementary ways to conduct fair and useful evaluations for effective programs?

Notes

1. Gordon Lippitt and Ronald Lippitt. (1978). *The consulting process in action.* La Jolla, CA : University Associates, Inc., p. 94.

2. Carol Weiss. (1972). *Evaluation research : Methods for assessing program effectiveness.* Englewood Cliffs, NJ : Prentice-Hall.

3. Lippitt & Lippitt, *op. cit.,* p. 40.

4. *Ibid.,* p. 36.

5. *Ibid.,* p. 33.

6. Weiss, *op. cit.,* pp. 104-107.

7. Lippitt & Lippitt, *op. cit.,* pp. 94-104.

The Author

Lyle F. Bachman is Associate Professor of English as an International Language and Director of Overseas Projects at the University of Illinois at Urbana-Champaign. He teaches graduate courses and has published widely in the areas of language testing, English for specific purposes, curriculum development and evaluation. His publications include *Reading English Discourse* (Prentice-Hall), and *Fundamental Considerations in Language Testing* (forthcoming, Addison-Wesley).