DOCUMENT RESUME

ED 214 988 TM 820 244

AUTHOR Fairbairn, Kerry
TITLE Fairness in Qualitative Evaluation: Some Implications for Educational Evaluators.
INSTITUTION Northwest Regional Educational Lab., Portland, Oreg.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
REPORT NO NWREL-REP-64
PUB DATE Oct 81
CONTRACT 400-80-0105
NOTE 28p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Educational Assessment; *Evaluative Thinking; *Evaluators
IDENTIFIERS Evaluation Problems; *Fairness; Objectivity; *Qualitative Analysis

ABSTRACT

The implications of the concept of fairness, to the level of the day-to-day business of the educational evaluator is explored. Section 2 presents a brief critical examination of the notion of objectivity in evaluation, and introduces the concept of fairness as an alternative focus for qualitative evaluation. In Section 3, there is a discussion of the ways in which the notion of fairness might be applied when an evaluation is being conducted. The issues addressed in this section include selecting the general style of the evaluation, formulating the questions to be answered, deciding on the sources of information, selecting methods for information collection, analyzing the information collected, and reporting the findings. The focus of Section 4 is the implications of adopting a "fairness" approach, for funding bodies and policy makers; for program participants and for evaluators. Some cost implications are also discussed. Finally, Section 5 draws together some conclusions about the use of a "fairness" criterion in educational evaluation. (Author/GK)

INTERIM DRAFT

Do not cite or quote without author's permission.

Author welcomes reactions and suggestions.

No. 64 FAIRNESS IN QUALITATIVE E'ALUATION: SOME IMPLICATIONS FOR EDUCATIONAL EVALUATORS

KERRY FAIRBAIRN

Queensland Department of Education

October 1981

Nick L. Smith, Director
Research on Evaluation Program
Northwest Regional Educational Laboratory
300 S.W. Sixth Avenue, Portland, Oregon 97204

Published by the Northwest Regional Educational Laboratory, a private nonprofit corporation. The work upon which this publication is based was performed pursuant to Contract No. 400-80-0105 of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.

PREFACE

The Research on Evaluation Program is a Northwest Regional Educational Laboratory project of research, development, testing, and training designed to create new evaluation methodologies for use in education. This document is one of a series of papers and reports produced by program staff, visiting scholars, adjunct scholars, and project collaborators--all members of a cooperative network of colleagues working on the development of new methodologies.

Are there instances when the fairness of an evaluation is more important than its objectivity? In what ways can evaluators improve the fairness of their studies? In this paper, Mr. Fairbairn answers these and related questions and highlights the implications of adopting a "fairness" approach to educational evaluation.

Nick L. Smith, Editor
Paper and Report Series

Contents

FAIRNESS IN QUALITATIVE EVALUATION: SOME IMPLICATIONS FOR EDUCATIONAL EVALUATORS

Section 1

Introduction

The literature on fairness in evaluation is small. A computer search across six of the well-known information bases unearthed nine references, only four of which address the topic in a substantial way. Of course there is a common-sense relationship between fairness and validity, and there exists a much larger literature on validity in evaluation.

However, although a relationship exists between validity and fairness, in the context of evaluation the notion of fairness provides a new source of ideas and implications for evaluators. The notion of fairness used in this paper goes beyond the usual understanding of validity, and encompasses such ideas as insuring that disadvantaged groups have their say, negotiating objectives and evaluation procedures with reference groups, and so on. An exploration of this notion of fairness might provide, for at least some evaluators, some new perspectives on the conduct of qualitative evaluations in educational settings.

The idea for this paper arose out of Egon Guba's (1978) paper on the possible value of using investigative journalism as a metaphor for educational evaluation. Guba highlights several notions and practices of investigative journalism that he judged to be of relevance to educational evaluation. Central among these is the notion of "fairness" employed in "interpretative" reporting. Guba contrasts this approach with that of evaluators

who have tended in the past to strive for objectivity. It is this notion of fairness, and not the methodology of investigative journalism, that forms the focus for this paper.

Ernest House (1980) presents a detailed analysis of concepts such as objectivity, subjectivity, justice and fairness. His overview should provide evaluators with a set of very useful guiding principles. House also discusses in some detail the need for and the possible elements of a fair evaluation contract. Every evaluator should have one!

The purpose of this paper is to further explore the implications of the concept of fairness, to the level of the day-to-day business of the educational evaluator. In a sense, it is one practitioner's viewpoint. I do not pretend to fully represent the views of either Guba or House on fairness in evaluation. Nevertheless, this paper owes much to their work.

Section 2 presents a brief critical examination of the notion of objectivity in evaluation, and introduces the concept of fairness as an alternative focus for qualitative evaluation.

In Section 3, there is a discussion of the ways in which the notion of fairness might be applied when an evaluation is being conducted. The issues addressed in this section include selecting the general style of the evaluation, formulating the questions to be answered, deciding on the sources of information, selecting methods for information collection, analyzing the information collected, and reporting the findings.

The focus of Section 4 is the implications of adopting a "fairness" approach, for funding bodies and policy makers, for program participants and for evaluators. Some cost implications also discussed.

lly, Section 5 draws together some conclusions about the "fairness" criterion in educational evaluation.

Section 2

Objectivity and Fairness in Evaluation

In the search for objectivity in evaluation, the following seem to be seen as important conditions:

- The evaluator should attempt to remain value neutral;
- There is one way of viewing a product or procedure, and the evaluator should strive to present this view;
- The opinions of individuals cannot be accepted unless those opinions are verified by many other individuals;
- Products and procedures can be compared by measuring them with common instruments that have been shown to be reliable (i.e., to produce replicable measures). These instruments should preferably use quantitative scales.

House (1980), Guba (1978), Zinkel (1979), Scriven (1972), and others have argued persuasively that the conditions necessary for objectivity do not exist in the real world. Each of the conditions listed above is open to argument.

- It is simply not possible for an evaluator to be value neutral. If he gives the impression of being value neutral, he will be seen as being uncaring. Rather, he must be impartial.

  > The impartiality of the evaluator must be seen as that of an actor in events, one who is responsive to the appropriate arguments but in whom the contending forces are balanced rather than nonexistent. (House, 1980, p. 92)[1]

- There are probably as many ways of viewing a product or a procedure as there are observers. Zinkel (1979, pp. 8-9) gives an example of a number of different people viewing an American flag. Not only do they view it from different physical positions, but each sees it in the light of his or her previous experiences. There is no single reality or meaning for the flag; it contains multiple realities.

- The quality of an opinion is not necessarily determined by the number of people giving that opinion. The notion of "objectivity" is a source of confusion here. House (1980) reports Scriven's analysis of objectivity into quantitative and qualitative objectivity. Quantitative objectivity is a sampling matter. If the "subjective" opinions of a number of individuals agree, one can say that one has represented the population. Qualitative objectivity, on the other hand, has to do with the quality of an observation, even if it is made by only one person. Groups are subject to biases as much as are individuals. It is just that the biases are different. There are occasions on which an individual's opinion may be of higher quality than that of a group.[2]

- In striving for replicability by using quantitative instruments to measure and compare products and procedures, there is the danger of measuring reliably the wrong things. That is, validity may have been sacrificed for the sake of objectivity (reliability).

So the notion of objectivity as the overriding criterion on which to build an evaluation is open to question. Some types of evaluations, namely those of a quantitative type, successfully meet some of the conditions of objectivity listed above. Within limits, then, objectivity-focused quantitative evaluations have their contribution to make. However, when the notion of objectivity is applied as the focus for qualitative evaluation, the objections to objectivity, as listed above, become crucial. I suspect that this is because qualitative evaluations tend to deal with complex issues and human interactions. My position in this paper is that qualitative evaluations should be grounded in the notion of fairness. This is not to say that objectivity is ruled out, but that the criterion of fairness would take precedence if there were a potential clash of interest between what is fair and what is objective.

Fairness, as I apply it to qualitative evaluation, is a complex, multi-faceted concept. Rather than attempt a succinct (and inevitably inadequate) definition of fairness, I have listed nine conditions that pertain to a fair evaluation.

1. A fair evaluation would be measured more by its balance (for example, in presenting differing views) than by its striving to unearth some immutable truth (Guba, 1978, p. 100).

2. The evaluator would get involved in the matter he/she is evaluating, so that he/she understands the range of viewpoints from an "inner-perspective".

3. The evaluator would attempt to represent faithfully all the views on the matter being evaluated, including opposing views. The evaluator would develop the ability to examine matters from multiple points of view, and would be prepared to report conflicting findings when they occur.

4. The reference group for an evaluation (the group from and about which information is collected) would be all those who would be affected by the program or policy being evaluated (House, 1980, p. 142).

5. The evaluation would be responsive in the sense that the evaluator would respond to what the participants (the various reference groups) see as the intended goals of their program or product and he/she would negotiate with them about appropriate ways of measuring these goals (House, 1980, p. 218).

6. The evaluator would ensure that the views of the least advantaged groups are given at least equal priority, if such groups can be identified (House, 1980, p. 134).

7. In selecting a methodology for a fair evaluation, the evaluator would try to ensure that there is no bias (class, sex, etc.) in the approach taken or in the methods used. Another sense in which the methodology should be fair is that it should provide an accurate index of the quality it is supposed to be measuring. That is, the methodology should be valid.

8. A fair evaluation would provide each audience with the type of explanation (usually, but not always, some kind of report), which is the best explanation for that particular audience. Explanations which make little or no sense to particular groups are not fair explanations (House, 1980, p. 89).

9. In terms of the evaluator's own position, a fair evaluator would

   - in fact do what he/she undertakes to do, or re-negotiate the agreement if necessary;

- go to lengths to rule out his/her own biases (Guba, 1978, p. 100);

- check the fairness of findings with participants and others at each opportunity.

The above list is certainly not exhaustive of the elements of a fair evaluation. If the list appears to be a self-righteous list of "shoulds", let me hasten to dispel this impression. Evaluations are usually complex things, and I feel the best an on-the-job evaluator can do is to strive to approach these and other conditions of fairness. If evaluators can move in the general direction of fairer evaluations, they will have made progress. But compromises will surely have to be made.

The purpose of the next section of the paper is to clarify some of these conditions of a fair evaluation by relating them to important decisions that all evaluators have to make both before and during the course of an evaluation. When it seems appropriate, I will draw on actual and hypothetical examples in an attempt to illustrate the points I am making.

Section 3: Applying the Notion of Fairness to Qualitative Evaluation in Education

As indicated earlier in this paper, Ernest House (1980) argues the need for fair evaluation agreements to be drawn up between evaluators, sponsors and other concerned parties. House draws on Care (1978) to provide a list of twelve conditions he feels are necessary for an evaluation agreement to be fair. These conditions are:

(i) Noncoercion. None of the participants in the agreement should be coerced or controlled by any of the others.

(ii) Rationality. All of the participants to the agreement should behave rationally.

(iii) Acceptance of terms. All parties to the agreement must accept the rules of procedure for reaching that agreement.

(iv) Joint agreement. The agreement should be reached after truly joint negotiations.

(v) Disinterestedness. No one party should pay excessive attention to its own interests.

(vi) Universality. The agreement arrived at should in a sense affect all parties to it equally. Each party would at least be prepared to occupy the position of any other party.

(vii) Community self-interestedness. The agreement reached should be in the best interests of the group(s) which will be affected.

(viii) Equal and full information. Participants should be equally informed of relevant facts.

(ix) Nonriskiness. The participants to the agreement should feel confident that every attempt will be made to implement it.

(x) Possibility. The agreement should be such that it is possible for the parties to carry it out.

(xi) Count all votes. All parties should be free to register whatever they wish to register in the final step of the agreement process.

(xii) Participation. All parties to the agreement should be allowed to have their say.

All evaluators have encountered situations during the course of an evaluation when they wished they had formalized an evaluation agreement. Uncertainties about release of information to evaluators, about data ownership, report circulation and editorial rights, etc., can make an evaluator's life very uncomfortable.[3] So there is no argument with House's proposition that a fair evaluation agreement is a necessary part of the evaluation procedure.

Evaluators live and work in many different contexts. Some of us are in positions to insist on some of the twelve conditions listed, but not on others. Stake and Easley (1978) in their Overview to the Case Studies in Science Education, refer to some of the constraints of contract research as it affected CSSE. For example, the National Science Foundation imposed requirements such as the number and representativeness of sites.

The important point that comes out of House's suggestions for forming fair evaluation agreements is that evaluators should try to the limits of their abilities, given their different work contexts, to achieve the fairest evaluation agreements possible. The repetition of this sort of striving is likely in the long run to improve the evaluator's chances of negotiating fair agreements, by changing the attitudes of other parties to the agreements, no matter how constraining the evaluator's work context was initially. Our aim in evaluation should be to move closer to the ideal, rather than to despair of ever achieving the ideal totally.

So the desirability of formulating a fair evaluation agreement along the lines outlined by House (1980) and Stake (1976) is not, I think, in doubt. But a fair evaluation agreement is only the beginning. I don't believe that a fair evaluation agreement guarantees a fair evaluation. It may be a necessary condition, but it is not sufficient.

Particularly if the evaluation is to be one which is responsive, which is negotiated continuously throughout its course, the evaluation agreement cannot spell out many of the details that could characterize the evaluation as fair or unfair. It may be necessary to gather information from sources and by methods not originally envisaged; it may even be necessary to change the basic questions being asked by the evaluator. I would therefore support Stake's (1976) proposition that an evaluation agreement should contain options for continuing negotiations.

There is a lot of room during the course of most evaluations for unfair practices to occur, regardless of the fairness of the initial contract. The rest of this paper is geared to what takes place after the negotiation of an evaluation agreement.

1. Formulating the Questions to be Answered

A fair evaluation would demand that the evaluation questions be formulated by the evaluator in consultation with the various reference groups for the evaluation. A program or a product is really a program or a product in operation, not a set of specifications on paper. There is no way to ensure that the right questions are being asked other than to consult the people who are involved in the operation. As House (1980) notes, the lack of consultation with sponsors in the Follow-Through Evaluation generated a set of unfair questions which the evaluation then attempted to answer.

The questions would not be the type that assume early closure. Rather, the questions would be worded in such a way that issues could be opened up and explored. The questions would assume the possibility of several answers, rather than one.

2. Selecting the General Style of the Evaluation

"Objectivity" may be sought in a rigid, pre-ordained evaluation design; "fairness" would not sit easily in such a design. Fairness would be more at home with responsiveness, changeability, because it is virtually impossible to anticipate, at the outset of a qualitative evaluation, all of the factors which might emerge, sometimes demanding new emphases, even new approaches. Larson and Kaplan (1981, p. 58) suggest the following reason for making the evaluation of programs adaptive rather than fixed.

> Information gathered during the course of an evaluation arrives in a probabilistic manner, providing evaluators with different knowledge profiles about the program at each intermediate point during the evaluation. Some, perhaps most, knowledge profiles should suggest a change in evaluation activities.

While it is possible that a cut-and-dried evaluation design may satisfy the demands of both objectivity and fairness given the right circumstances, the chances of this happening are fairly remote.

An evaluation design that is able to encompass the emergence of new issues, the appearance of new data sources and the ongoing feedback of reference groups is much more likely to be characterized by balance, by presenting multiple perspectives and by the kind of understanding that will provide for each audience an adequate explanation of whatever is being evaluated.

While the Follow-Through Evaluation did change during its history, the changes were from one fairly rigid design to another, none of which seemed able adequately to represent the programs in operation. And the changes in design were brought about by "external" factors such as changes in direction handed down by the Office of Education, and by changes in the contracting groups, rather than by meaningful negotiations between evaluators and reference groups.

An evaluation style that is negotiated, open, responsive and flexible will likely produce a qualitative evaluation that is fair. Given these conditions, elaborate pre-emptive designs would be much less likely to produce fair evaluations than designs that have modest, tentative beginnings and grow throughout the course of the evaluation.

3. Deciding on the Sources of Information

Earlier in this paper it was suggested that the reference groups for an evaluation should be those groups that are (or are to be) affected by the program or product being evaluated. This means that information for the evaluation should be collected from each of these groups. It does not mean that the same information should be collected from each of them by the same information collection methods. Rather, if useful information is to be collected from each group, it should be collected in ways that are agreed to by the group and are meaningful to the group.

Time constraints and shortage of money seem to be constant companions of educational evaluations. Consider the predicament of the evaluator of a new curriculum unit in, say, consumer education, who has very limited time and not very much money to

complete the task. He/she has reviewed the committee decisions and other documentation that led to the new unit, and has conducted some type of context analysis on the curriculum unit itself. The evaluator has yet to gather the reactions of reference groups, but has identified teachers, students, parents and local business organizations as the groups most likely to be affected by the new curriculum unit. It would be good if it were possible to gather the opinions of sufficient members from each of these reference groups so that he had a "reliable" sample of each. However, time and money constraints make this an impossibility. The decision to be made, then, might be between these two alternatives:

(a) Limit the number of groups from which to collect information to, say, two groups, but select a "reliable" sample from each group; or

(b) Collect information from a smaller and technically unreliable, sample of the members of each reference group.

By selecting alternative (b), the evaluator would have chosen to come closer to satisfying the requirements of a fair evaluation than if he had chosen alternative (a). This is not to say that the evaluator does not value reliability, but that, forced to choose, he opts for fairness in the belief that the evaluation thus will be of more use to those concerned.

4. Selecting Methods for Information Collection

If it is accepted that fairness is predicated on, among other things, the recognition of multiple realities and multiple perspectives, it follows that fair evaluations will need to employ multiple methods to reflect this notion. Generally speaking, evaluations that use a variety of information collection methods for each information source, and that seek information from a variety of sources (as in (3) above), will have satisfied a number of the important conditions of fairness outlined in Section 2 of this paper, providing the methods have been negotiated with the groups supplying the information.

In negotiating with reference groups about information collection methods, the question of appropriateness is important. A fair evaluator would be aware of biases that could be introduced if certain methods were used with certain groups. The culture bias of some IQ and achievement tests is well known; needs assessment surveys can be class-biased (House, 1973); written surveys can be inappropriate for less literate groups; for some national groups, any form of survey can represent a frightening bureaucratic intrusion, and so on. Again, it is in the evaluators' best interests to talk to reference groups about ways in which they can best provide information for the evaluation.

Methods that postpone closure about issues have a better fit with the fair evaluation approach than do methods which attempt rapidly to reach a single conclusion. The recognition that the complexities of qualitative evaluations require subtler interpretations, should logically discourage the evaluator from seeking quickly arrived-at, single, authoritative conclusions. As House (1980) notes, adversary evaluations, which model themselves on courtroom procedures, attempt to force an authoritative conclusion, whereas some of the case study approaches postpone such conclusions, leaving room for enlarged discussion and more subtle interpretation. The evaluator may be required, in the end, to recommend one course of action over all other possible courses of action. In a fair qualitative evaluation, the decision on which this recommendation is based would be delayed until alternative perspectives had been thoroughly explored within the limits of the operating constraints with which the evaluator is working.

5. Analyzing the Information Collected

In most qualitative evaluations, a major task of information analysis is to reduce the mass of information collected to manageable proportions. Therefore, decisions have to be made about what information to use, what information to condense, and

how to analyze it. In making these decisions the evaluator has a good deal of control over the fairness of the evaluation.

It is tempting to reduce the data to those that readily fit into some existing analysis format, such as a statistical package. Inevitably, some data will not fit into such an analysis format, without being transformed into something bearing little resemblance to the original. Yielding to such a temptation at the information analysis stage can render an otherwise fair evaluation unfair with a single blow. So a fair evaluation would strive to analyze each type of information in ways that best preserve the character of the original. Perhaps this approach to analysis might not yield results that are readily "comparable", but if the original information had been somehow violated to achieve comparability across information types, the comparability would be very superficial anyway.

Negotiation is as important at this stage in a qualitative evaluation as it is at other stages. Putting the information through a "black box" for processing not only brings into question the evaluation's credibility, it also detracts from its fairness in fact. Just as reference groups can provide important inputs about how information might best be collected from them, they can also provide useful perspectives on how this information might best be analyzed to preserve its meanings. So to neglect this stage of an evaluation is to overlook important information.

6. Reporting the Findings

If the evaluation has been a fair one, the reported findings will usually present few major surprises to the reference groups and others centrally involved in the evaluation, because they will all have been consulted to some degree about the findings as they became available. And to some degree, the reported findings will have been molded by the reactions of the reference groups to the tentative findings presented to them during the evaluation. So there will probably be little "new" information as such.

But the question of how the information is reported is crucial to the evaluation's fairness. House (1980) says that

> unless an evaluation provides an explanation for a particular audience, and enhances the understanding of that audience by the content and form of the arguments it presents, it is not an adequate evaluation for that audience, even though the facts on which it is based are verifiable by other procedures (p. 89).

A quick reading of a random selection of evaluation reports will show that many of them (probably the majority) seem to have been written for other evaluators, rather than for the evaluation's reference groups and others interested in the evaluation. It appears sometimes as if educational evaluators are trying to communicate to other educational evaluators on the assumption that

> one can communicate statements to another scientist which are reducible to what the other scientist will experience if he does certain things (House, 1980, p. 58).

This view would assume that "objectivity" is achievable and all-important in evaluation. The thesis of this paper is that context is multi-faceted and crucial in evaluation; the validity of findings must therefore be demonstrated. This can usually be done in the type of language most readily understood by the audience(s). Often this will be everyday, non-technical language.

There has not been much research on the effects of evaluation reports of various kinds on various audiences. It cannot be guaranteed that a "fair" evaluation report will be fairly received. Because in a fair evaluation the reader may be explicitly required to play a quite active valuing role, it is important that the information provided in the report be a balanced account.

Section 4: Some Implications of Adopting a Fairness Approach

Adopting fairness as a major criterion for qualitative evaluations would cause some changes in the ways evaluators, program participants, policy-makers, and funding bodies act during the evaluation, and in the ways they are affected by the evaluation. Some of these changes are outlined below.

Evaluators

Fair qualitative evaluations require that evaluators understand different views about a program from an inner perspective. Thus more emphasis would be put on methods involving interaction with participants, and less emphasis would be put on impersonal approaches. Special qualities of interpersonal sensitivity and rapport would be required of the evaluator in order that he/she be able to understand and impartially report conflicting views. Questionnaire surveys and other "impersonal" methods will probably always remain part of many qualitative evaluations. However, there would need to be greater emphasis, in fair qualitative evaluations, on participant observation methods, interviewing, and other "personal" or direct methods of information collection. An evaluator would certainly need to have a wide range of methods available in order to tap multiple information sources in a variety of ways.

Not only would the evaluator's methods be diverse and more "personal", but they would also tend towards the kind that do not force rapid closure. That is, the evaluator would increasingly use methods that delay the forming of conclusions. Case study methods would certainly play an important role in many "fair" evaluations.

There would also be implications for the evaluator in the overall design of the evaluation. It would be very difficult to come up with an elegant evaluation design and stick to it throughout the evaluation. A fair evaluation would require that

the evaluator be open and responsive to changes in contextual factors, to the emergence of new issues and to changes in the nature of program aims. An evaluator who is concerned about being "fair" would therefore be one who could live with an evolving evaluation design, who felt easy with a lack of "structure", at least in the early stages of the evaluation.

Program Participants

The major implications for program participants would have to do with their time commitment to the evaluation, which would probably increase somewhat, and with the style of their contribution, which would change from being reactive to being more interactive. Participants would enter into a closer relationship with the evaluator as he/she tried to understand their views, motives and values; they would be asked to validate the evaluator's perceptions of their views, and might also be asked to review tentative conclusions from the evaluation.

It is likely too that some of the participants in the evaluation would take part in the drafting of an evaluation contract, and in ongoing negotiations about the direction and nature of the evaluation. Participants would have more control over the ways in which their programs are to be evaluated. They would be parties to the evaluation in a real sense.

In exchange for the increased commitment of time and energy, participants in a fair evaluation should feel that their views really count, that they have made a positive contribution to the evaluation, and perhaps understand their own values and motives a little better as a result. The alienation that participants often feel on reading the final evaluation report should, therefore, be less common in a fair evaluation.

Policy-Makers and Funding Bodies

Policy-makers, funding bodies and others who commission evaluations and respond to the results of evaluations would also play a changed role in a "fair" evaluation. They may be

involved, for example, in key negotiations about the nature and conditions of the evaluation, leading to the formulation of an evaluation contract. The contract would be very much a negotiated document, not just a set of specifications and requirements handed down by a funding organization or a policy-maker.

Beyond the initial contract there may well be further negotiations about the nature of the evaluation, as one or another of the parties to the evaluation perceives the need for a change in direction. Therefore, the role of the policy-maker or funder is likely to be more active, and to continue throughout the evaluation.

Because reports of evaluations concerned with fairness as a central issue are less likely to present single answers in the form of clear-cut recommendations, policy-makers and funding bodies are likely to have to take a more active role at the end of the evaluation as well. They will be asked to consider a number of arguments, some of them possibly contradictory, and form conclusions based on these arguments. However, as with program participants, program funders and policy-makers, having been relatively closely involved with the evaluation, should not be too surprised by anything in the evaluation report. Indeed, they may well be able to predict what will be in it.

Above all, a certain amount of trust will be asked of funders and policy-makers who contract to have a "fair" evaluation conducted; it may be difficult for them at first to have this trust if they are used to contracting for evaluations based largely on the criterion of objectivity. After all, in a "fair" evaluation they would be relinquishing some of the tight control traditionally held by them, in favor of negotiating the course of the evaluation more or less on equal terms with evaluators and reference groups.

Costs

There are so many variables involved, that it is difficult to estimate the costs of fair qualitative evaluations relative to others that are focused on objectivity. However, it is possible to speculate on some aspects of relative costs and benefits.

We could probably afford, in many circumstances, to sacrifice some of the "accuracy" of the more traditional quantitative evaluation approaches, and improve fairness characteristics without spending any extra money. If both fairness and objectivity were to be improved, the overall costs would most likely be higher.

There is a potential extra benefit of fair evaluations that ought not to be overlooked. The higher level of participation by program participants and funding bodies in the evaluation process should give them greater insights into the program or product being evaluated and into evaluation itelf. That is, there is a potential training effect that, in both the short term and the long term, could be valuable. Of course there is also an associated cost in time for the greater involvement of program participants and funding bodies.

Section 5: Concluding Comments

The processes of a "fair" evaluation are likely to expose many difficult issues that are often not even encountered in evaluations (for example, the existence of very different understandings of seemingly simple concepts), because they remain unseen, but not because they are not there. So evaluating with fairness is not likely to be easy. However, the resulting evaluations should be better for having faced up to these difficult issues.

In practice, evaluators will have to approach the question of fairness with a deal of common sense. A quote from Anderson and Benjaminson (1976, p. 163), who were writing about investigative reporting, illustrates this point:

> A factually accurate story, however, should never be killed because it is "unfair" in the sense that contrary assertions are not published along with the damaging facts. The stories published about the space program, for example, were true even though they didn't contain rebuttals from the Flat Earth Society.

These authors go a little too far, though, in urging investigative reporters not to be pre-occupied with trying to achieve fairness, and putting the responsibility for unearthing bias squarely on the shoulders of the reader, thus:

> a knowledgeable reader will realize it [bias] and the reporter's work will accomplish very little beyond impressing his or her relatives (p. 161).

I do not think unfairness or bias is often as easy to identify as Anderson and Benjaminson would have us believe, even for the "knowledgeable" reader. Therefore I feel that the responsibility for fairness in evaluations should rest with us, the evaluators.

The responsibility is a fairly serious one. House (1980, p. 170) says that fair procedures in the formulation stages do not guarantee a fair evaluation agreement. I would extend that thought and contend that fair procedures alone, during the conduct of an evaluation, do not of themselves guarantee a fair evaluation. In the final analysis, the evaluator's own impartiality comes into question. In House's words, "He [the evaluator] must give evidence of his impartiality by showing how he has acted contrary to his own interests in the past." (1980, p. 93). And of course there is no absolute guarantee of impartiality to be inferred from that statement either. In the end, the matters discussed above should make for fairer evaluations; they can not guarantee them.

Two thoughts occurred to me on reading a draft of this paper. First, I have been writing about "the evaluation" in some sort of encapsulated sense, as if it had a separate existence. Ideally, evaluation would blur into program experience and might not be very obvious as a separate entity. Second, I have set up a rather artificial contrast between evaluations "that are and used to be", and "fair" evaluations of the future. Of course

many evaluations have captured what I have called "fairness" long before the term was ever used in the literature. The extreme contrasts presented as a vehicle for the main idea of this paper are probably very rare in reality.

NOTES

1. Zinkel (1979, p. 18) advocates that the inquirer adopt a position of "complete neutrality" in order to establish the "essential nature of the reality". Others, such as House, would deny the possibility of neutrality and argue that it is for the audience (of the evaluation report) to establish the nature of the reality.

2. The social psychologist Irving Janis has warned of a conformity phenomenon, "groupthink", which can occur when a dominating leader has a group gathered around him. Through collective rationalization, groupthinkers ignore warnings that might otherwise lead them to reconsider their assumptions. Janis gives a hypothetical example of a group of Presidential advisers reacting to the question: How do we deal with the Communist threat? Each, as a member of the group, wants to be at least as extreme as the others, and preferably a little more so. If a person shows signs of reticence, he's isolated as a softie. (Reported in Omni magazine, May 1981).

3. Helen Simons (1981) points out that the evaluator's role becomes especially difficult when dealing with "superordinates"--executives and policy managers who perceive their boundary control and image management to be undermined by the evaluation. While she sees the need for formal evaluation contracts, Simons doubts that even a formal contract could withstand the type of pressure that executives can mount to withhold information, limit the distribution of reports, and so on.

References

Anderson, D. and Benjaminson, P. Investigative reporting. Bloomington & London: Indiana University Press, 1976.

Case, N. S. Participation and policy. Ethics, July, 1978, 88.

Guba, Egon G. Metaphor adaptation report: Investigative journalism. Research on Evaluation Program Paper and Report Series. Portland: Northwest Regional Educational Laboratory, 1978.

House, E. R. The conscience of educational evaluation. In E. R. House (Ed.) School evaluation. Berkeley: McCutchan, 1973.

House, E. R. Evaluating with validity. Beverly Hills, London: SAGE Publications, 1980.

House, E. R., Glass, G. V, McLean, L. D. and Walker, D. F. No simple answer: Critique of the follow-through evaluation. Harvard Educational Review, 1978, 48(2).

Larson, R. C. and Kaplan, E. H. Decision-oriented approaches to program evaluation. New Directions for Program Evaluation, 1981, (10).

Leff, Walli F. Groupthink. Omni Magazine, May 1981.

Scriven, M. Objectivity and subjectivity in educational research. In L. G. Thomas (Ed.) Philosophical redirection of educational research. National Society for the Study of Education, 1972.

Simons, Helen. Case studying superordinates: The power politics of policy evaluation. Paper presented to a Symposium on Case study in policy evaluation: Paradoxes of popularity, Annual Meeting of the American Educational Research Association, Los Angeles, 1981.

Stake, R. E. Evaluating educational programs: The need and the response. Washington, D.C.: OECD Publications Center, 1976.

Stake, R. E. and Easley, J. A. Case studies in science education. Urbana-Champaign: Center for Instructional Research and Curriculum Evaluation, 1978.

Zinkel, Cal. Phenomenology, multiple realities and educational evaluation. Research on Evaluation Program Paper and Report Series. Portland: Northwest Regional Educational Laboratory, 1979.