

EXPLORATIONS IN
EVALUATION THEORY DEVELOPMENT

Marvin C. Alkin, Editor

CSE Report No. 123

December 1978

Center for the Study of Evaluation
Graduate School of Education
University of California, Los Angeles

Table of Contents

	Page
Part I. Introduction	1
1. An Approach to Evaluation Theory Development	3
Marvin C. Alkin	
Part II. Role Playing and Response	
2. My Views on Evaluation after having read <u>The New Rhetoric</u> -- by "Robert Rippey".	9
Marcella Pitts	
3. Effects of <u>The New Rhetoric</u>	15
Robert Rippey	
4. My Views on Evaluation after having read <u>The New Rhetoric</u> -- by "Egon Guba".	23
Elaine Lindheim	
5. The New Rhetoric of Egon Guba Compared to that of Ten Years Ago.	27
Egon Guba	
6. My Views on Educational Evaluation after having read <u>The New Rhetoric</u> -- by "Robert Wolf"	31
Richard Daillak	
7. <u>The New Rhetoric: A Case for the Role of Clarification in Educational Evaluation</u>	37
Robert Wolf	
Part III. Theory Development	
8. Evaluation Theory Building	43
Ernest House	
9. An Inquiry into the Nature of Evaluation Theorizing	49
Marvin C. Alkin and Frederick Ellett	
References.	63

Introduction

The basic theme for this issue is the question of evaluation theorizing and the nature of theory building. The initial papers emerged from a graduate class exercise in which students read The New Rhetoric by Ch. Perelman and L. Olbrechts-Tyteca and then used this treatise as a framework for a role-playing exercise.

In these papers, the actual authors (Marcella Pitts, Elaine Lindheim, Richard Daillak) "assumed the role" of a specific evaluation theorist (Robert Rippey, Egon Guba, Robert Wolf). These papers were entitled, "My Views on Evaluation after Having Read The New Rhetoric." Use of the first person in these papers resides solely with the actual authors and the approval (of Rippey, Guba, Wolf) herein should not be inferred.

At the 1978 Annual Meeting of the American Education Research Association, the role-playing papers provided the initial basis for a symposium session followed by the responses to these papers by the actual theorists. And, a commentary by Ernest House and an introduction paper by Marvin Alkin completed the presentations at the symposium session. Subsequent to the meeting, the tape transcriptions of the presentations were edited by Professor Alkin and revised, and in some cases substantially modified by the theorists. A philosophical assessment of the nature of evaluation theory based on the symposium was written as the final paper in this collection by Alkin and Frederick Ellett.

AN APPROACH TO EVALUATION THEORY DEVELOPMENT

Marvin C. Alkin

Graduate School of Education

University of California, Los Angeles

Any theory is a theory about the basic data of the set of phenomena that are being examined; and those basic data have to be clear. The exercise reported in this symposium was designed to help in structuring some of the preliminary data necessary for evaluation theory development. And just what is the nature of the basic data to which I refer?

The basic data are the intricacies, characteristics, and descriptions of the various entities referred to as evaluation "theories." A potential evaluation theory must, in fact, be a theory which comes to grips with the various theories of evaluation. That is, an evaluation theory must include within it the various evaluation "theories"--let us refer to these instead as kinds of evaluation. Thus, one function of the exercise we have engaged in (and are continuing today) is the expansion and refinement of data related to three kinds of evaluation (the Egon Guba kind, the Robert Rippey kind, and the Robert Wolf kind).

Now, one way to expand and refine data related to different kinds of evaluation is to devise a situation which demands that each of the evaluation kinds be defended both on rational and empirical grounds. Possibly the most straightforward way of attaining this end is to simply question the developer of a particular kind of evaluation and, as a necessary first step, seek clarification about the different points of view expressed that

appear not to be clear. That procedure for attaining clarification and then expansion of evaluation types has occurred many times over the years at the various AERA meetings; and the expansions and clarifications ensuing have tended to be rather superficial in nature.

A preferable approach would be, perhaps, to have Egon Guba available for a period of three or four weeks, with no other demands upon his time, to fully discuss the dimensions of his various views on evaluation. But this approach to developing the data for evaluation theory building also has its failings. Often the question and answer approaches with primary attention and focus on a theoretician's own writings can lead to a restricted defense and/or expansion of the author's point of view. Focusing on the writings related to the kind of evaluation tends to restrict the questions that might be asked or the clarifications sought.

Greater insight into an evaluation type and a better understanding of the evaluation conceptions of individuals can be gained by forcing them to scrutinize their viewpoints in light of a different kind of conceptual framework presented as an input. In the process of responding to an external conceptual framework, an author will be provoked and challenged and hopefully forced to expand and clarify and to defend his viewpoints. Thus, while the data related to a particular kind of evaluation can be expanded by having that type of evaluation explicitly described, there is apparently a higher level of data explication; that higher level is the forced application of the evaluation kind to a situation, or the examination of the implication of a kind of evaluation to a specific known stimulus such as an external conceptualization.

device to gain insights into the nature of different evaluation kinds. A unique study, set within the context of a doctoral level evaluation seminar, was designed. Each of the participants in the seminar selected a prominent evaluation theoretician and spent the initial weeks of the course becoming thoroughly familiar with that theoretician's works. There needed to be a procedure to insure that each of the participants had acquired an appropriate level of familiarity with "their" theorist. Thus, a discussion meeting was scheduled modeled upon Steve Allen's American educational television talk show, "Meeting of the Minds." This allowed each participant the opportunity to act out the part of his/her chosen theoretician by making a presentation of his views and interacting with his colleagues there assembled.

Following this class exercise, the participants spent six weeks systematically reading and discussing in class The New Rhetoric. At the conclusion of the examination of this treatise, participants were asked to write papers... "indicating modifications to 'your' views (role playing your theoretician) after having read The New Rhetoric." These papers were subsequently discussed, edited, refined, and presented at an American Educational Research Association symposium. This symposium offered a further refinement on the theory development methodology; additional clarification and expansion related to each of the evaluation types took place by having Guba respond to "Guba," Rippey to "Rippey," and Wolf to "Wolf."

And now, before turning to comments from the Gubas, Rippeys, and Wolfs, it is appropriate to present a brief overview of the nature of

the external conceptualization used as a stimulus in this evaluation theory exercise. A dominant mode in Western philosophy for the last three centuries has been set by the Cartesian concept of reason and reasoning. In contrast, Perelman and Olbrechts-Tyteca (1969) have put forth a treatise devoted to a consideration of argumentation, a concept set in the ancient tradition of Greek rhetoric and dialectic--"the art of persuading and convincing, the technique of deliberation and discussion."

However, their point of view is at some variance with the ancient rhetoric, in that Greek rhetoric was concerned primarily with persuasive public speaking. Perelman and Olbrechts-Tyteca, on the other hand, are concerned with written argumentation. Nonetheless, many of the notions from traditional rhetoric are applicable to the new rhetoric as well. One such area of ancient, as well as modern day, concern is the idea of "audience." Every discourse, every bit of rhetoric, is intended for a particular audience. Is this not an idea which has great applicability to evaluation as well? Of the 105 sections of this book, seven deal with the question of audience, including the concepts of the "universal audience" and the "audience as a construction of the speaker," to name two of the less esoteric-sounding headings.

Major topics addressed in the Perelman and Olbrechts-Tyteca book include: the framework of argumentation, the starting point of argument, agreement, the choice of data and adaptation for argumentative purposes, presentation of data and form of the discourse, quasi-logical arguments, arguments based on the structure of reality, the relations establishing the structure of reality, the disassociation of concepts, and the interaction of arguments.

Without being able to provide even an inkling of the full complexity and depth of this treatise, I am sure that the reader will agree that it would certainly seem to be an excellent and meaty external conceptualization to be used as a stimulus for further refinement of several kinds of evaluation.

My Views on Evaluation after having
Read THE NEW RHETORIC -- by "Robert Rippey"

Marcella Pitts

Graduate School of Education
University of California, Los Angeles

We would all agree that evaluation theories develop over time, gradually evolving through a series of successive changes. This evolution is influenced by several sources of input, including practical experience gained in conducting evaluations, discussions with colleagues, and extensive reading both within and outside of education. While all these sources inspire the evaluation theorist, the influence of readings outside of education is of particular interest since these "unexpected influences" can bring a new and refreshing perspective to developing evaluation theories.

Recently I have encountered such an unexpected influence in the form of The New Rhetoric, a philosophical treatise on argumentation which has provided me with new insights into my own theory of evaluation, transactional evaluation. I would like to share with you my thoughts on transactional evaluation after having read TNR. First, I will briefly define transactional evaluation and discuss some basic premises of argumentation as they are presented in this philosophical work. Then I will discuss what I see as basic similarities between transactional evaluation and argumentation. Finally, I will share with you some thoughts I have on the role the transactional evaluator plays in light of the premises set forth in TNR.

Transactional evaluation has been developed as an alternative to the more conventional forms of evaluation which focus exclusively on the outcomes of a program and thereby fail to examine the problems and conflicts its implementation generates in the system undergoing change and in the changers themselves. The problems and conflicts they are experiencing are the target of transactional evaluation for I believe that the failure of many programs is due to the inability of institutions and their personnel to respond to the demands for change brought about by a new program. They resist change and their resistance is often expressed as hostility, conflict, procrastination, or subversion of the program. Transactional evaluation attempts to deal with the conflict generated by change that can undermine a new program. By doing this, transactional evaluation creates a climate for change. It analyzes the dysfunctions caused by threats to stable roles and clarifies these roles. It identifies and attempts to reduce any resistance to change on the part of individuals who might be affected by the program by encouraging them to participate in the evaluation. The program is implemented on an experimental basis and there is continuous evaluation of its consequences by those who support the program as well as by those who do not support it. In the course of the transactional evaluation process these opponents and proponents confront each other and negotiate the conflicts which might prevent the successful implementation of the program. Basically, transactional evaluation is an innovative approach to evaluation which is designed to bring about the adoption of change.

I am struck by the parallels that exist between the transactional evaluation process and the process of argumentation as it is described in TNR.

Argumentation is also employed to bring about change--change in people's belief in and commitment to an idea. It is used when absolute proof or conclusive evidence, which in and of itself, would cause people to endorse an idea, is not available or cannot be obtained. In the absence of such "proof," argumentation seeks to induce or increase people's allegiance to a point of view or a course of action. A successful argument either causes people to act in a desired way or exhibit a willingness to act in a desired way. For argumentation to exist, certain conditions must be met. One of the most fundamental of these is agreement, in principle, on the formation of an intellectual community to debate a specific question together (Perelman & Olbrechts-Tyteca, 1969, p. 14). The parties must establish intellectual contact and must agree on the basic issues on which deliberation is to begin. The nature of the transactional process which draws together divergent parties and establishes a relationship between them creates the community of minds that is needed if argumentation is to take place.

In addition to creating conditions which are conducive to argumentation, transactional evaluation is related to argumentation on another level. The goals of transactional evaluation and those of persuasion, the type of argumentation which it has been observed elsewhere as applicable to evaluation (House, 1977), are compatible. Persuasion is successful if it increases the adherence of a particular audience to a thesis which is presented for their assent (Perelman & Olbrechts-Tyteca, 1969, p. 29). Transactional evaluation is successful if antagonists of the program are persuaded to support it or, at very least, its transactional evaluation.

As a part of transactional evaluation both proponents and opponents of a program work together to develop and implement or have implemented a transactional evaluation plan. ~~The inclusion of critics of the program~~ in this planning effort gives them a legitimate and constructive role to play, a role which can lead to their conversion to support of the program and to the utilization of their skills and ideas to modify and improve the program. The transactional evaluation plan they help design is developed as the program develops, rather than before the program gets underway as is the case with traditional formative and summative evaluation plans. It seems to me that this kind of planning is akin to the practical approach which can be employed to resolve incompatibilities in argumentation. A practical approach resolves problems as they arise. When applied to evaluation, it creates a flexibility which makes transactional evaluation the ideal tool to use with politically sensitive programs. It permits the rethinking of rules and concepts in terms of the constantly changing demands, conflicts, and decisions which must be made in a particular evaluation setting.

Throughout the transactional evaluation process the evaluator has a central role to play. First, the responsible evaluator will see to it that the program does not fail for lack of insight into factors that may be impeding its implementation. Secondly, the transactional evaluator, as I have come to realize after reading TNR, by creating a situation that is conducive to persuasive argumentation, is an integral part of the argumentative process. This is a somewhat problematic position to be in. Ideally, the transactional evaluator acts as a catalyst by initiating the process of

confrontation and negotiation between conflicting interest groups and helping persuasive argumentation proceed in a more conciliatory atmosphere than might otherwise be possible. The evaluator would not influence the outcomes of the argumentative process nor actively participate in it. Standing for a balance of forces, paying maximum attention to the interests at issue, dividing his or her attention equally among the different points of view, the evaluator would remain impartial. In the capacity of a facilitator, the evaluator would offer clarification of the issues involved in the evaluation context, draw out people's concerns, find a common ground for the program's proponents and opponents to begin deliberation, and provide continuous feedback to concerned parties. What I have gleaned in TNR about argumentation, however, makes me pause and wonder about the role of the transactional evaluator. This role, as I have just described it, seems to lend itself very well to purposeful as well as inadvertent argumentation on the part of the transactional evaluator. The implications of this for transactional evaluation are considerable. The evaluator would be an active participant in the argumentative process and in that capacity would bring a different kind of influence to bear on the outcome of the transactional evaluation process. I am not adverse to this possible shift in the role of the transactional evaluator, although it seems to indicate that we in the field of evaluation may have to reconceptualize our views of what an evaluator can and should do in the name of evaluation.

In closing I would like to say that TNR was a thought provoking book for this evaluation theorist. I believe that a later paper in this symposium will also discuss the possible impact of TNR on the theory of transactional evaluation. I look forward to that presentation.

EFFECTS OF THE NEW RHETORIC

Robert M. Rippey

University of Connecticut

Preparing for this symposium, I summarized my thoughts on evaluation before reading The New Rhetoric (Perelman & Olbrechts-Tyteca, 1969). First, I made the classical distinction between measurement, evaluation, and assessment. Evaluation had something to do with reaching objectives. I divided objectives into two categories: individual objectives and program objectives. I then restricted my scope to program evaluation and to exemplify my conceptual base, outlined a product--a four-part evaluation report. Part one of the report would be a program description, a program design, and a rationale for expecting the design to accomplish the objectives.

Part two would be the methods section. I feel that descriptions of methods are of great importance following Joe Schwab's dictum that knowledge has little meaning without a description of how it is arrived at. No knowledge is certain and certitude only occupies some problematic region within the bounds of the validity of available evidence.

Part three would report on implementation. What was done? Where did the program depart from design and why? What were the impediments to implementation?

The fourth part of the evaluation report would indicate the effects of the program on the students, the teachers, and the institution. This looked rather pedestrian. Perhaps it would stay that way.

In addition to this summative evaluation report, formative evaluation

would occupy the evaluator in an active role, pointing out which objectives were in difficulty and also (this is the transactional part) what difficulties were being experienced by personnel as a result of the change: who was having trouble meeting his or her new role expectations and why, and what might be done about it.

The idea of the evaluator, in the name of objectivity, sitting on his can while a project goes down the chute does not appeal to me. But doesn't the activist evaluator affect outcomes? I should hope so. After all, what does "formative evaluation" mean? This leads to the conclusion: Replication of studies of intensively evaluated innovations should include a similarly active evaluator until the absence of evaluator/treatment interaction can be demonstrated.

The active, home-based evaluator is the kind of person and role we were encouraging back in the sixties at the University of Chicago Center for the Cooperative Study of Instruction: someone with lots of technical savvy, but someone who was where the action was as well.

After reading TNR, which I liked intensely, I looked for changes in my ideas and behavior. This is what I observed. My first rhetorical act and change in behavior was a visit to and presentation to the Glastonbury Board of Education. I opposed their policy of lowering students' grades when they were absent and not allowing an appeal until their grades reached the "F" level. I discovered the board was in general agreement with me and it was the teachers who were strong on the policy so I followed Perelman's advice of addressing a specific audience and rewrote the talk for a public hearing the following week. In both presentations, I used double hierarchies,

utilized dissociation by means of philosophical pairs, and presented my arguments in Nestorian order. We won.

~~-----~~ I also prepared and presented at a departmental seminar, a list of 102 words I had never heard of before I read the book. (See Figure 1.) Now we can all write evaluation reports which absolutely no one will understand. The value of this ability is pointed out by the motto of my class of staff associates at the University of Chicago who acknowledged for dissertations, and by analogy evaluation reports, "to be precise is to be found out." I also found the concept of epideictic discourse useful in preparing funeral orations for educational pseudo-innovations which had died either natural or violent deaths.

More seriously, I really appreciated Perelman and Olbrechts-Tyteca's delineation of the differences among impartiality, objectivity, partisanship, fanaticism, and skepticism. I have always lacked patience with evaluators who let projects sink in order to maintain objectivity. Is indifference a prerequisite for objectivity? TNR says, "In spheres where thought and action are closely mingled, impartiality stands between the objectivity which fails to qualify the third person for interference, and the partisan spirit which absolutely disqualifies him." What they are saying is that the evaluator should not be a voyeur. If he is not a member of the team, he will not be able to speak with authority because commitment and intimacy will be lacking. The objective evaluator is as much disqualified for his remoteness as the project director for his partisan bias. Perelman and Olbrechts-Tyteca go on to say, "Whenever it is necessary to refute the accusation that our desires have determined our beliefs, it is essential that we

Figure 1

List of Terms I Never Heard of Before or Which I Thought I Knew and Didn't

- | | |
|---------------------------------------|------------------------------|
| 1. Univocity | 48. Distortion |
| 2. Qualifiers | 49. Allusion |
| 3. Notions | 50. Personification |
| 4. Loci | 51. Apostrophe |
| 5. Presence | 52. Prosopopoeia |
| 6. Starting points | 53. Participation |
| 7. Facts | 54. The symbolic relation |
| 8. Truths | 55. Metonymy |
| 9. Presumptions | 56. Synecdoche |
| 10. Periphrasis | 57. Hierarchies |
| 11. Word families | 58. Double hierarchies |
| 12. Petitio principii | 59. Amplification |
| 13. Ad hominem | 60. Arguments of fortiori |
| 14. Ad rem | 61. Arguments of degree |
| 15. Ad humanitatem | 62. Arguments of order |
| 16. Ad personam | 63. Quantitative |
| 17. Prolepsis | 64. Qualitative |
| 18. Hypotactic discourse | 65. Sorities |
| 19. Paratactic discourse | 66. Structure of reality |
| 20. Modalities: | 67. Example |
| Assertive | 68. Particular to particular |
| Injunctive | 69. Exemplum in contrarium |
| Interrogative | 70. Exception |
| Optative | 71. Illustration |
| Argumentative use of pronouns, etc. | 72. Antiphrasis |
| 21. Quasi logical arguments | 73. Anti-model |
| 22. Autophagia | 74. The perfect being |
| 23. Retort | 75. Analogy |
| 24. Argumentative role of definitions | 76. Theme |
| 25. Role of justice | 77. Phoros |
| 26. Arguments of reciprocity | 78. Metaphor |
| 27. Arguments of comparison | 79. Trope |
| 28. Arguments of transitivity | 80. Dormant metaphors |
| 29. Arguments based on probabilities | 81. Dissociation of concepts |
| 30. Arguments based on structure of: | 82. Philosophical pairs |
| Reality | 83. Persuasive Definition |
| Causes | 84. Anantiosis |
| Pragmatics | 85. Otiose |
| Ends and means | 86. Hesitation |
| 31. Direction | 87. Hyperbaton |
| 32. Stages | 88. Inversion |
| 33. Contagion | 89. Parataxis |
| 34. Popularization | 90. Asteism |
| 35. Consolidation | 91. Pseudo-converse |
| 36. Unlimited development | 92. Corax |
| 37. Hyperbole | 93. Amplitude |
| 38. Litotes | 94. Reticence |
| 39. Irony | 95. Renunciation |
| 40. Interaction of act and person | 96. Self-renunciation |
| 41. Arguments from authority | 97. Concession |
| 42. Severence | 98. Epitrope |
| 43. Restraint | 99. Exordium |
| 44. Surrender | 100. Nestorian Order |
| 45. Act and essence | 101. Anticipatory refutation |
| 46. Abuse | 102. Prolepsis |
| 47. Deficiency | |

Figure 1

List of Terms I Never Heard of Before or Which I Thought I Knew and Didn't

1. Univocality
2. Qualifiers
3. Notions
4. Loci
5. Presence
6. Starting points
7. Facts
8. Truths
9. Presumptions
10. Periphrasis
11. Word families
12. Petitio principii
13. Ad hominem
14. Ad rem
15. Ad humanitatem
16. Ad personam
17. Prolepsis
18. Hypotactic discourse
19. Paratactic discourse
20. Modalities:
 - Assertive
 - Injunctive
 - Interrogative
 - Optative
 - Argumentative use of pronouns, etc.
21. Quasi logical arguments
22. Autophagia
23. Retort
24. Argumentative role of definitions
25. Role of justice
26. Arguments of reciprocity
27. Arguments of comparison
28. Arguments of transitivity
29. Arguments based on probabilities
30. Arguments based on structure of:
 - Reality
 - Causes
 - Pragmatics
 - Ends and means
31. Direction
32. Stages
33. Contagion
34. Popularization
35. Consolidation
36. Unlimited development
37. Hyperbole
38. Litotes
39. Irony
40. Interaction of act and person
41. Arguments from authority
42. Severence
43. Restraint
44. Surrender
45. Act and essence
46. Abuse
47. Deficiency
48. Distortion
49. Allusion
50. Personification
51. Apostrophe
52. Prosopopoeia
53. Participation
54. The symbolic relation
55. Metonymy
56. Synecdoche
57. Hierarchies
58. Double hierarchies
59. Amplification
60. Arguments of fortiori
61. Arguments of degree
62. Arguments of order
63. Quantitative
64. Qualitative
65. Sorities
66. Structure of reality
67. Example
68. Particular to particular
69. Exemplum in contrarium
70. Exception
71. Illustration
72. Antiphrasis
73. Anti-model
74. The perfect being
75. Analogy
76. Theme
77. Phoros
78. Metaphor
79. Trope
80. Dormant metaphors
81. Dissociation of concepts
82. Philosophical pairs
83. Persuasive Definition
84. Anantiosis
85. Otiose
86. Hesitation
87. Hyperbaton
88. Inversion
89. Parataxis
90. Asteism
91. Pseudo-converse
92. Corax
93. Amplitude
94. Reticence
95. Renunciation
96. Self-renunciation
97. Concession
98. Epitrope
99. Exordium
100. Nestorian Order
101. Anticipatory refutation
102. Prolepsis

furnish proof, not of our objectivity, which is not possible, but of our impartiality." TNR provides one method of demonstrating impartiality. The methods of science do likewise by means of design, randomization, and double blinds. Let me give you examples of impartiality taken from a evaluations of two projects to which I was deeply committed.

Project #1 was aimed at improving the writing skills of students in grades nine and ten. One year after the completion of the program, students who had had the program and students who had not were randomly mixed in classes taught by teachers different from those involved in the project. These teachers in grade eleven were asked to assign an essay topic to grade using their regular grading standard, whatever it might be. There was a significant difference in these grades favoring the experimental group. Tell me where my bias was showing on that one.

Project #2 involved frequent classroom meetings among parents, teachers, and first and second grade children to talk about how to get along in school and how to facilitate communication among adults and children. I visited third grade classes, again where participants and non-participants were mixed and where participants and non-participants were not known by the teacher. The project director randomly paired six children in each of a number of classrooms--three project children and three non-project children. I spent an hour in each class observing behaviors at one minute intervals switching from child to child. I didn't know who had been in the project. I noted such things as attentiveness, daydreaming, conversing with an adult, bothering other children, etc. Again I got a significant difference in project objective behaviors favoring project participants and no significant differences among behaviors not emphasized in project objectives.

One of the most profound concepts I found in The New Rhetoric was the necessity of establishing a community of minds prior to argumentation. I have seen evaluation reports misused as often as not, sometimes to turn off an excellent project and other times to prolong the life of a dead duck. Perhaps the evaluator, who is usually in the position of advocate whether he likes it or not, needs to spend more time establishing a community of minds committed to argumentation before he begins his work. I have previously advocated the use of project monitoring committees made up of both the protagonists and the antagonists of proposed innovations. These committees would advise and participate in the formative, summative, and transactional evaluation of the innovation. There are probably some antagonists and also some protagonists who should not be on this kind of committee; TNR clarifies who these should be: those who will not join the community of minds. I quote, "Recourse to argumentation assumes the establishment of a community of minds which, while it lasts, excludes the use of violence. To agree to discussion means 1) readiness to see things from the viewpoint of the interlocutor, and 2) to restrict oneself to what he admits and to give effect to one's own beliefs only to the extent that the person one is trying to persuade is willing to give assent to them."

I am anxious to explore the effect of the establishment of such a community of minds, prior to evaluation. Perhaps there is no point in proceeding before this has been done. I have always asked, "What is the purpose of the evaluation? What will its results be used for?" but I have not sought an adequate commitment to argumentation from all hands.

I struggled reading and trying to understand TNR. I reverberate now that my battle with their many esoteric words is past. What I have just said is what I can say today. Sensing the continuing reverberations, who knows what the effects will be tomorrow?

Figure 2

Addendum March 2

Dear Diary:

I just received a copy of House's (1977) Logic of Evaluative Argument today. I read it immediately and appreciated his synthesis of Perelman and Olbrechts-Tyteca (1969). The contents were old hat to me --a diligent scholar of Perelman and Olbrechts-Tyteca but I wished I had started with House's work. It would have made the path easier. But the struggle was worth it. House compels me, however, to make additional comment about an influence of TNR which I had deliberately concealed in my January 30 draft. Eve had three faces and I have at least that many and one of my faces has to do with something called confidence testing. I knew that Marcella had looked at my stuff on transactional evaluation. She even sent me a letter asking if I had any other material and what I sent her was restricted to transactional evaluation. However, after reading House's monograph, I feel the truth must out. On page 7, House contrasts evaluation as argumentation with evaluation as persuasion. He contrasts credibility with certainty, compulsion with necessity, variable adherence with truth and falsity, ambiguity with clarity, etc. Confidence testing seems to me, and seemed to me even on my first run through TNR, to be a suitable methodology for studying evaluation as argumentation. It involves the use of probabilities, not single choices, in responding to test items (either essay or multiple choice) and through complex methods of scoring and analysis, divides a person's score into a component representing certain knowledge and a component representing perceived knowledge. High perceived knowledge in the absence of certain knowledge sounds like adherence to me. Applied to test items having unique correct answers, and applied to items where answers are ill-defined, even to the experts, this method makes possible the quantitative study of the adherence of minds. I have recently completed the development of the necessary technology and have performed some preliminary validity studies which are promising. Perelman and Olbrechts-Tyteca provided a stronger argumentative base and additional encouragement for my work in this direction.

MY VIEWS ON EVALUATION AFTER HAVING READ THE NEW RHETORIC

-- by "Egon Guba"

Elaine Lindheim

Graduate School of Education

University of California, Los Angeles

In reading The New Rhetoric, I found myself agreeing with the premise of the authors that argumentation can be used to bring about an understanding of how and why decisions are made. I also found that many of my basic beliefs about what evaluation is and how it should be done were reinforced. Finally, I came away with new insights and possible strategies for approaching some of the evaluation utilization problems which have been concerning me recently. (See, for instance, Guba, 1975.) In fact, I might even propose that a good way to evaluate the book itself might be to implement some of the arguments it discusses in an evaluation report, and then compare how effectively that evaluation is utilized as versus a similar evaluation not structured around argumentation. In other words, can following TNR make a significant difference in how well an evaluator serves a decision maker?

My basic definition of evaluation has not changed since reading this book. Evaluation is the process of obtaining and providing useful information for making educational decisions (Guba & Horvat, 1970). For a client to be well-served, the information provided by an evaluator needs to be timely, credible, relevant, and above all, usable. The premise of the authors of TNR that argumentation is aimed at bringing about

an understanding of how and why decisions are made, as well as at bringing about some actual future action or change in disposition, therefore has great appeal for me. If using rhetorical principles makes it possible to persuade others to action, then an evaluator who masters those principles will be able to present information in a compelling manner.

My evaluation model focuses upon the decision maker as the central determinant in why and how any information is gathered. (See Guba & Horvat, 1970.) Therefore I read with special interest those portions of TNR which dealt with styles of decision making or problem solving. One section detailed three quite different approaches to dealing with conflicts or incompatibilities: the logical approach, the practical approach, and the diplomatic approach. I think that if an evaluator could get to know a decision maker well enough to form a notion of which style that individual favors, evaluation results could then be presented to that individual in the manner best suited to his or her style. This approach might also at least partially allay some of the concerns I expressed in 1969, about our lack of knowledge about decision-processes (Guba, 1969).

Another concern of mine has been that different interest groups may hold different purposes for the same evaluation. (Guba, 1975 detailed seven factors influencing a client's problems in evaluation utilization; this was the first factor.) Often these purposes are more covert than communicated, meaning that the evaluator must do some careful probing in order to pinpoint what the foci of the evaluation really are. It is necessary to determine whether there really are decisions to be made, and if so, whether intrinsic or extrinsic criteria will be used in

resolving those decisions. Quite possibly, the evaluator will find that his activities are only a front, serving to ratify, legitimize, or vindicate decisions already made. Therefore, the dialectic technique detailed in TNR might be a useful tool for an evaluator to use when first conceptualizing the evaluation. Via questions and answers, the points which are to be decided and the proofs which will be acceptable and conclusive ones can be determined. An evaluator who could derive such agreement points from a decision maker would have removed one more obstacle in the way of effective utilization of the information to be presented.

Closely related to my concern about hidden interests is my awareness that the audiences for which an evaluation must be prepared often will differ in their values, concerns, and interest, as well as in the stake they have in the evaluation results. (This is the fourth factor discussed in Guba, 1975.) Again, the TNR seems to offer some practical strategies, in this instance through the concept of audience presented by the authors. The book's repeated emphasis on the need for a speaker to know well those whom he wishes to influence by argumentation would hold true for evaluation as well. Often an evaluator will need to persuade a composite audience and to employ a multiplicity of arguments. Remembering that argumentation must begin with the views of the audience, and not with any personal sentiments on the part of the speaker, seems crucial. Indeed, the same information might be presented via entirely different arguments to two different audiences, simply because the arguments appropriate for one would be ludicrous for the other.

Several of my most major concerns with the utilization of evaluation

THE NEW RHETORIC OF EGON GUBA COMPARED TO THAT OF TEN YEARS AGO

Egon G. Guba

Indiana University

As you know, an important member of the international Community Party is that functionary known as the "Party theoretician," whose function it is to rationalize whatever new directions the Party leadership wishes to take while all the while making it appear that the new moves are entirely consistent with the Party's historical commitments. Revisionism is Communism's cardinal sin; the theoretician's task is to make all the changes seem to be non-revisionist. I find myself, having read what Elaine Lindheim had to say, trying to make the present state of my mind appear to be perfectly consistent with my former position (which Elaine has captured so well), even though I know it is not.

For you see, in many ways I am not the "real" Egon Guba--at least not the one that Lindheim had in mind when she conducted her exercise. That Egon Guba was the one working and writing seven to 10 years ago. If, as she said, she had some problems in understanding what that Egon Guba was like, I presently share her dilemma; my mind has changed about so many things that it is hard for me to recall what opinions I held then. I would, for example, surely not, today, try to construct a model of evaluation based on decision-making processes, although ten years ago that seemed to be exactly the way to go.

It is always difficult for the developer of a theory to know, at any given stage, just where his mind is--what assumptions he is making, what

data he is arbitrarily including and excluding, which opinions he takes into account and which he eschews. As Ernest House (1977) has pointed out in his splendid The Logic of Evaluative Argument, it is really impossible to "prove" anything in an absolute sense; in the final analysis, evaluators, and indeed, all inquirers, can only hope to persuade others to their point of view, not to compel them by the sheer weight of irrefutable evidence. The theory builder is also constructing a persuasive argument, based on axioms and postulates that are certainly arguable. The kind of exercise in which we are engaged helps both builder and critics become clearer as to just what the substance of the argument is, and provides an opportunity to test the degree of its persuasability.

Now Lindheim's analysis has helped me to see where I was ten years ago. If I had then read The New Rhetoric I might very well have been affected by it in the ways she has suggested. I think she has done a fine job of sensing where I was then. But my problem this morning is not to validate her statements but to suggest what the real Egon Guba of 1978 would say on reading TNR. And that is really quite different from what Lindheim has suggested.

I certainly would not find in TNR today corroboration for a definition of evaluation as a process of servicing decisions. In fact, I would find in it some basis for suggesting that the decision model (at least insofar as decision making is thought to be some kind of rational process in the sense defined by Braybrooke and Lindblom, 1963) is not appropriate for the evaluator at all. It seems to me now to be a mistake to imagine that one can construct a great matrix whose rows are defined by decision options and whose

columns are defined by criteria, and that, by entering appropriate data in the cells, one can assess each option against all criteria and so come to a judgment of the best option. Such a formulation is based on assumptions which, although not clear to me in 1967, are now evident and now found by me to be nonpersuasive.

What are some of these untenable assumptions? Let me list some of them: that there is a decision authority (decision authority is in fact diffuse and hard to pin down); that decisions are made in some explicit way (decisions more often simply "bubble up" and when they are officially recognized, e.g., in some ratification vote, have already been "made" in all important operational senses); that there is a decision time (there is almost never an exact decision time--decisions can be made and remade, postponed and pushed ahead, etc., ad infinitum); that there is a set of criteria commonly agreed to and based upon a common value structure that is shared by the important others in the situation (values are often multiple and in conflict). Thus the notion of a logical, rational, synoptic (to use Braybrooke and Lindblom's 1963 term) decision process is hardly defensible.

I am much more inclined now to think along the lines suggested by House in his (1977) monograph. There are, he asserts, various audiences to an evaluation persuadable by different kinds of evaluative argument. The task of the evaluator is to construct a rhetoric that is appropriate to the variety of audiences that exist. But that objective opens a can of worms for the evaluator. What constitutes ethical behavior in that context? Does ethical behavior include slanting data for different audiences, picking out what they want to hear and telling them only that? Or does it include representing to

each of the audiences what the perspectives and values of other audiences are? Does it mean working with the audiences so that they will achieve consensus? Is the notion of consensus even appropriate? I doubt it, judging from my own experience.

Well, these are the kinds of considerations that would be paramount in my mind were I to read Rhetoric today. I am sure that I would find corroboration for that view of evaluation rather than for a model of evaluation based on decision processes of the sort that I was talking about ten years ago.

I must give you an A for effort, Elaine Lindheim, and an A for reconstructing me as I probably was a decade ago. In fact, you may have done a better job than I myself am able to do in that reconstruction, for you are probably able to take a more objective view than I. But that reconstruction does not reflect accurately my 1978 posture. You of course have no way of knowing about that new posture from a reading of my published work. Little of my present position has found its way into literature.

I hope that these changes came as something of a surprise to you, but not totally so. It should not, after all, be so surprising that my thinking has changed a bit over 10 years. Consistency, it is said, is the hobgoblin of small minds. We all reserve the right to change our minds. Hopefully, in this case, it is because my thinking has matured.

My Views on Educational Evaluation After
Having Read THE NEW RHETORIC -- by "Robert Wolf"

Richard H. Dailak

Center for the Study of Evaluation

Graduate School of Education

University of California, Los Angeles

Although the title of this paper indicates that I will today discuss my views on educational evaluation, it seems appropriate to begin with a few remarks on the book, TNR (Perelman & Olbrechts-Tyteca, 1969), which stimulated this paper and our discussion here. TNR is a philosophical treatise on "argumentation," which we might (simply) define as the process by which we persuade others, thereby gaining their adherence to some point of view or set of beliefs. Perelman and Olbrechts-Tyteca note that the study of argumentation has, since about the time of Descartes, been slighted, even ignored, while scholars have concentrated instead on formal logic and mathematical or quasi-mathematical demonstration and proof. They contend that the exclusion of argumentation is "a perfectly unjustified and unwarranted limitation of the domain of action of our faculty of reasoning and proving" (p.3); reasonable persons do attempt to persuade others, even when they lack "irrefutable proof" of the correctness of their contentions, and the reasonable, rational listener can be persuaded by convincing argument.

Thus, Perelman and Olbrechts-Tyteca believe that argumentation is a legitimate activity of rational human beings, and the study of argumentation should be renewed. TNR is a first step in this direction;

in it, the authors analyze the methods of argumentation and argumentative "proof" employed in the human sciences, law and philosophy and develop a quite rigorous classification and analysis of the structure and technique of human argument.

There are numerous applications one might make of the ideas elucidated in TNR. My colleagues at this symposium have discussed a few of the insights on evaluation that can be derived from the book; many other useful applications could be found. In this paper, however, I hope to take a step back from concrete application in order to consider more the overall thrust of the book and its relationship to an approach to evaluation -- one emphasizing "naturalistic/judicial inquiry" -- which I have advocated elsewhere (Wolf, 1973, 1975; Wolf & Tymitz, 1977a, b).

Briefly, Naturalistic Inquiry and Judicial Evaluation are two compatible, and complementary, methodologies for investigating policy issues, developing a base of information relevant to the issues under study, and presenting that information in a manner which clarifies the issues while preserving the integrity and coherence of differing points of view. A major strength of naturalistic/judicial inquiry is its ability to deal with "subjective" data, especially people's beliefs, concerns, and attitudes, within a framework which clarifies the alternative interpretations people hold and guides the observer in judging the merits of these interpretations. As Tymitz and I have described it:

Techniques of naturalistic inquiry [such as the intensive probing, open-ended interview] allow for an in-depth investigation of the issues as they are perceived by numerous persons who function within a variety of roles and settings. Procedures within JEM [the Judicial Evaluation Model] allow for the public presentation of the organized data in what we have identified as a Clarification Forum. The JEM relies on the law's acceptance of human

testimony to clarify and subsequently judge complex events. Facts, perceptions and opinions are exposed, cross-checked and expanded upon by the investigators and the respondents themselves. Concepts from both jury trials and administrative hearings in the field of law offer a useful system of evidentiary rules and procedures aimed at producing alternative inferences from data prior to the rendering of policy recommendations. (Wolf & Tymitz, 1977a, p. 4)

Clearly, the fundamental thesis of TNR, that argumentation, deliberation, and reasoned judgment are legitimate and important, is compatible with the perspective implicit in the naturalistic/judicial inquiry approach. In discussing the utility of the judicial evaluation model, I have often commented on the need to include human judgment as a source of evidence in evaluation and policy decision making (e.g., Wolf, 1973, 1975), and I have criticized the tendency in the evaluation community to rely too heavily upon the dominant social science research paradigm, which emphasizes quantification and complex analysis. Thus, it is heartening to read the persuasive introductory remarks of Perelman and Olbrechts-Tyteca regarding the validity and utility of precisely the kinds of evidence most valued in naturalistic/judicial inquiry, and it is gratifying to note their support of argumentation and deliberation, two of the key components of the judicial approach.

TNR does, however, suffer certain limitations which may restrict its usefulness to the evaluation theorist. First and foremost is its emphasis on categorizing and analyzing the techniques of argumentation; almost the entirety of the book is taken up with this task. Very little attention is given to the process of deliberating upon arguments made by others, and the book offers virtually no guidance to the reader who may be interested in honing his or her skills at weighing the arguments advanced by a partisan advocate of this or that position.

Thus, TNR focuses on a skill which can be an important component of a system for constructively exploring controversial issues, but which could be used manipulatively and unscrupulously if detached from the constraints of the surrounding system. It is therefore distressing that Perelman and Olbrechts-Tyteca discuss the techniques of argumentation at such length, yet give the reader so little guidance either in privately judging the merits of conflicting arguments or, better still, in devising systems, like the judicial system, which harness argumentation to productive ends.

Much of my own work in evaluation theory has centered around developing systematic approaches to controversial issues -- such as the Naturalistic Inquiry/Judicial Evaluation Model -- so the reader should not be surprised at my conclusion: TNR is agreeable in its valuing of argumentation, deliberation, and judgment, but it is a long way from furnishing solutions to the real problems of evaluation, its amoral treatment of argumentation is somewhat disconcerting, and its total influence on my views has been slight.

In fact, some of our most recent experiences in the application of naturalistic/judicial inquiry (Wolf & Tymitz, 1977a) are especially supportive of this conclusion. In applying naturalistic/judicial inquiry to a policy study of Public Law 94-142, we found the pre-argumentation phases, including the framing of the issues and the collection of quality data on these issues, to be critically important and, perhaps, in need of refinement; TNR has little to offer in this matter. And the experience of others (e.g., Popham & Carlson, 1977) is testimony to the need for well designed constraints on argumentation so that the judicial evaluation paradigm does not become simply adversary debate.

THE NEW RHETORIC:
A CASE FOR THE ROLE OF CLARIFICATION
IN EDUCATIONAL EVALUATION

Robert L. Wolf

Indiana University

My task is to discuss The New Rhetoric (Perelman & Olbrecht-Tyteca, 1969) in terms of the Judicial Evaluation Model (JEM), and comment on the theme of the work in relation to the underlying aspects of judicial evaluation procedures. While it is my belief that TNR offers a useful example of reasoning processes that depart from the traditional statistical inference paradigm, the work leaves much to be desired in terms of guiding evaluators toward a better consideration of evidence, deliberation, and clarification. In fact, TNR can even be hurtful to impartial evaluation efforts, because of its particular relevance to partisan advocacy.

In essence, the most salient criticism that I would direct toward TNR and particularly House's (1977) interpretation of it, is the emphasis on argumentation without a clear articulation of how argumentative skills can benefit the field of educational evaluation. Sometimes we mislead ourselves in believing that we, as evaluators or social scientists, do not engage in argumentation anyway. I believe that much of what we do in trying to confirm our hypotheses and convince others that our data are right (accurate) exemplifies argument and persuasion. So, I do not believe TNR is all that new to the way we actually conduct our affairs as evaluators. Many times we act as defenders of a point of view, trying to convince and persuade others that our data are the best data, or our interpretation of them should prevail. The Judicial

Evaluation Model (JEM) that I have been working on for the past several years departs from this unitary style of persuasion. The JEM recognizes human fallibility, and assumes that evaluators have biases, and that they view the world from different frames of reference. These points of view not only influence the collection and analysis of evidence, but even more fundamentally, the kinds of questions we pose and ultimately the manner in which we present our findings. The JEM allows for the juxtaposition of those different viewpoints and different interpretations of evidence. Argumentation and persuasion become illuminating rather than obfuscating.

In creating a conceptual structure for conducting evaluation studies it is essential that a cross-checking or cross-validation procedure be required. Persons who benefit from the outcomes of evaluation would benefit more if they could examine different inferences, or additional explanations that can be generated by any information display. That such procedures involve the presentation of opposing points of view, structured around particular issues and substantiated by different evidence, keeps the inquiry reasonable and impartial. Furthermore, these sorts of cross-validation processes enhance credibility, and provide a more comprehensive assessment of the data than more conventional evaluation strategies. Of course, the former demands data collection processes which facilitate in-depth understanding. In this regard, for the past three years I have continued to refine the operationalization of natural inquiry strategies which demonstrate powerful applicability to an elaboration of inferences and explanations. The essence of good evaluation, therefore, should not rely upon the skills of argumentation because we already employ those kinds of skills, often to the disservice of clients or program consumers. More useful skills would be learning how to understand

the significance of contrasting arguments and how to challenge or extend inferences and explanations. I contend that any set of opposing arguments presented in some systematic fashion, with a clear set of rules that guide such presentation, helps to clarify some of the complexities, nuances, and subtleties of educational endeavors that simply do not get clarified when only one person is arguing his point of view or only one evaluator is presenting his/her findings. Deleting or de-emphasizing this clarification aspect of contesting arguments is a serious limitation to TNR and to its theoretical utility.

It should be noted, however, that while clarification is an important part of evaluation it is not an end in and of itself. The clarification process must be structured in such a manner as to facilitate and enhance the quality of decisions that are made, the recommendations that are offered, or the changes that take place. Again, the JEM is an approach that places a premium on clarification, but it does so with the expectation that clarification will lead to some meaningful action, and therefore, provides a structure for that action to occur. Obviously, part of that structure involves the presentation of good sound arguments which will hopefully provoke a range of potential actions or help generate possible strategies for change and modification. The ability to argue points of contention in a clear and meaningful way will increase the likelihood that clarification will result and thus increase the probability that the evaluation process itself will promote action.

Within the JEM framework, TNR attention to argumentative technique is useful. Being able to construct rational and intelligible arguments,

and present those arguments clearly are essential skills for evaluators employing judicial evaluation procedures. But it would be an egregious oversight to simply develop those skills without understanding the context for which those skills need to be mastered. And, the context, as I see it, is a forum where evaluators can present opposing arguments based upon different evidence or different interpretations of the evidence that is relevant to the issue(s) under consideration. Without such a presentation and without striving for clarification, argument will simply result in persuasion. And persuasion only serves the side that wins.

So critical is this latter point that it warrants illustration. In a recent article in the Educational Researcher (Popham, 1977) a critique of adversary evaluation (which some people often confuse with judicial evaluation) dramatized the dangerous abuse of argumentation and persuasion in evaluation. Popham reported a study of a curriculum project in Hawaii where two evaluation teams debated different interpretations of the same set of data. The evaluators relied upon rhetoric and persuasion. There were no witnesses, no cross-examination, no intent to structure a genuine clarification process. In short, this was not judicial evaluation, but rather a misuse of the idea that allowed persuasion to prevail. And that idea, simply stated, concerns the use of sound, reasonable, and contrastive presentations for the purpose of clarifying program complexity so that informed decisions can be rendered and significant changes made.

In the final analysis, if we cannot strive for such clarification then the underlying rationale for evaluation is questionable. Similarly, if we continue to conduct studies that are biased by one point of view, i.e., one

argument, or excessively rely on persuasion, then those studies will continue to be unfair--much like the Hawaii experiment was.

I believe that TNR points to the need for better argumentative skills, but it is important to remember that argument for argument's sake is not a worthwhile evaluation goal. If, on the other hand, those skills can be employed within a structured clarification process, such as the JEM affords, then I believe that the development of such skills is an important step in the right direction. For any process that helps keep the many sides of truth alive is worth nurturing and admiring.

EVALUATION THEORY BUILDING

Ernest House

University of Illinois

Theory building - how new theory construction comes about - is the import of this exercise, so I will address my comments to theory building rather than the ideas of my logic monograph or the Perelman book.

I first saw some notice of the Perelman book about 1970-71. I was reading the Key Reporter, which is a quarterly newsletter of Phi Beta Kappa. There was a one or two sentence review of the book. They often review many books--20 or 30 in one issue. The review said that it was a translation from the French which attempted to deal with modes of non-scientific discourse. That rang a bell with me. So, I sent away for the book which was about \$20.00 (that was expensive at that time). After I received the book, I put it aside for several years until about 1974 or 1975, when I started trying to read it. It was kind of accidental that I started to read the Perelman book at that time and yet in a sense - that event is not accidental. My reading of The New Rhetoric was not simply an adventitious incident.

At that time in evaluation, I was looking for a different way of getting at what programs were all about. I put some mention of non-scientific discourse, as a matter of fact, into a paper I wrote entitled "The Conscience of Educational Evaluation" (House, 1972). Bob Wolf was a graduate student at CIRCE* concurrently and picked up on that idea. Wolf was already interested in the legal aspect and took off developing a lot of the adversary ideas. The point I want to make is that all this occurs within a long tradition.

*Center for Instructional Research and Curriculum Evaluation, University of Illinois, Champaign-Urbana.

There are those who say that evaluation has made no progress over the last ten years. But, I think that is not the case. It is a tradition, a discourse, a mode of thought that goes back at least 10 years, to the mid 1960's, and certainly Egon Guba has been a very important part of that.

When I started the gifted evaluation at Illinois, I set up an advisory panel and with Egon Guba, Dan Stufflebeam, and Bob Stake on it. Many of the ideas that I started working on then were ideas derived from those people at that time. I see that actually what has happened is that it is a continuous interplay of ideas. In a sense it is a dialectical development, as it were, of the thoughts involved in evaluation. So, although Guba was renouncing his earlier self of years ago, I feel differently. In fact, that earlier self, the Egon Guba of 10 years ago, was a very important part of what contemporary thinking about evaluation is now. That is, we evaluation theorists have played off his ideas (and those of others) to get to the ideas that we now have. We really have a tradition here - an intellectual tradition that we are trying to develop.

I also should say, that my background is one very deep in the humanities so that it is not accidental that I should pick this kind of stuff to write about since it actually goes back to and even prior to, Aristotle. So, I am versed in that kind of humanities having background in Philosophy, literature, and the like. I was very susceptible to seeing evaluation as argumentation, as discourse. I could see scientific discourse as being one form of argument within a larger universe of discourses. In education, what we have tried to do is be pseudo-scientific, with not very good results for the most part. Guba was saying this about education ten years ago. We had kind of run out of gas in the line of the development we were pursuing.

Another thing you need in such a theory development is a knowledge of what has gone before in evaluation. Having worked in this field for ten years, I am heavily steeped in all these evaluation studies that we have done and sometimes I feel burdened by them all. As a theorist, you have to know, in a sense, what has happened in the field in order to relate the theory - the abstract ideas - to it. I sometimes see the educational philosophers or others trying to develop theory for field-like evaluation or for field-like education that often does not really come off because they do not understand the elements with which they are working. That is, if you are going to develop a theory in evaluation you had better know what is really going on in evaluation. You have to know the studies and you have to know the tradition of the people who are in the field. Knowing that kind of thing, I set about trying to embody the Perelman ideas and other ideas, and to put them into a form that we would find recognizable in the evaluation field. I have cited lots of works in evaluation. I went to the works of some of the saints in the field - Cronbach and Campbell - and looked at their writings knowing that they have been moving in this direction. I used their publications as kind of a jump-off into some of the Perelman reasoning.

I have also used examples. I took an evaluation example which Gene Glass did in the Educational Researcher about six years ago (Glass, 1972). It was an analysis of some evaluation that Scriven had made for the American Educational Research Association. I took Glass' evaluation and analyzed his evaluation in terms of the reasoning that he employed using Perelman's categories to show that Glass was actually using certain kinds of reasoning that Perelman says one might use as opposed to what we often think as scientific

discourse. Scriven (1972) fortunately responded to Glass and I was able to analyze Scriven's reasoning too. It was interesting, (see House, 1977), that neither of them used data to any great extent. But, I think data is greatly over-rated in evaluation. I don't see this work as simply being a set of techniques that one might employ. That is not the case at all. It is a misuse if you say "O. K., I am going to use means end reasoning here." It is a little bit like discovering that you are speaking prose. That is, you are already using this kind of reasoning. Evaluators already use this kind of reasoning. What my monograph does, and what Perelman does, is say that indeed these are legitimate modes of reasoning. We are already employing these ideas; we must be more conscious of what we are doing - more critical and more reflective of what we are doing. We are really talking about different ways of knowing other than the scientific way of knowing. It is really epistemological problem that we are dealing with.

Finally, I'd like to make a comment with reference to Guba's point about theory rationalization and the theorizers, rationalizers. I am afraid that I must confess to doing a little bit of each. Because, we already are doing many things in evaluation that we have no real justification for. I felt rather deliberately that I wanted to construct some sort of rationalization to justify many of the things that we are already doing. Within the frame of reference that I have, within a different way of knowing, one doesn't go necessarily from theory to practice. That is, one sees knowledge as being a working back and forth between the practice of evaluation and the theory of evaluation. One affects the other. It doesn't go merely one way.

It is alright to try to justify what in some cases we are already doing in practice. It is not a hypocritical position. The mode of knowing is much more a dialectical mode of knowing.

Finally, I think this whole exercise that Alkin has constructed is pretty ingenious from the pedagogical point of view. And, it is an interesting attempt at evaluation theory development and refinement. It is too bad that my students don't get this kind of instruction and teaching. They will have to go to UCLA to get this kind of imaginative stuff. It is a very interesting way to conduct a class. Have you evaluated this, Marv?

AN INQUIRY INTO THE NATURE OF EVALUATION THEORIZING

Marvin C. Alkin

Frederick Ellett

University of California, Los Angeles

This paper is about evaluation theorizing. Our focus is the various strategies that people follow when they are engaged in theorizing -- that is, in developing evaluation theory. We are not interested in criticizing the conclusions they have reached. Nor are we interested in expanding and redefining their evaluation theories. Instead, we are trying to focus on the various strategies that people might use in developing their theory.

There are several benefits to be gained from examining the nature of theorizing. One of these is a better understanding of the nature of evaluation theory development. By understanding the process better one is able to engage in theory construction more reflectively and carefully. Secondly, an understanding of evaluation theory development will allow one to better understand the nature of evaluation theory generally and of the various evaluation theories. For example, when someone is developing a theory, many of the conjectures are very tentative. Even though the person's statements may give others the impression that a certain issue is settled, the real status of that issue may, in fact, still be tentative. The tentativeness is usually lost to readers who may not think of the issue in that way. Thus, we believe that engaging in an examination of the evaluation theorizing process should be useful to students who are interested in understanding theory as it evolves. Practitioners might come to be more critical of

theoretical pronouncements. It will give theoreticians a more realistic understanding of the nature of the enterprise.

The way we have chosen to begin the process of examining the methodological considerations in the development of evaluation theory is by considering the personal insights afforded by each of the individual papers within our small set. The ultimate goal and the goal that we are working toward is to come to understand the more informal and personal aspects of theory development. We shall use this small set of cases as a first step in explicating these informal and personal aspects. It is to be understood that we don't pretend to be comprehensive or conclusive in our work here. We hope the reader and especially the theoretician will now have a general understanding of the ultimate goal so that he might join us in its pursuit.

By the personal, more informal aspects we mean to mark off the following distinctions. There is some tendency to regard theory development and particular thinking as something akin to geometry -- as an exact mathematical procedure. What we want to emphasize is the essential role of the more personal and informal aspects of theory development. When people are engaged in theory development they follow certain strategies and they come to grips with certain aspects or states of theory development which have been largely ignored because they had thought of theory development only in a deductive mathematical sense. We want to highlight these aspects. For example, it is our belief that by studying these cases in a more intimate, more naturalistic way we will be able to draw out principles - sound principles useful in theory development.

Our reasoning for including the students' cases in this set of papers

is twofold. First of all, the use of the student papers (Pitts, Lindheim, Daillak) illustrates a very interesting pedagogical technique. (A discussion of these aspects are not within the scope of this paper.) Our second reason for having the student papers included -- having students reconstruct the views of Rippey, Guba and Wolf -- provides the occasion for the authors to reconsider and restructure their own thoughts. Papers by the students portraying or reconstructing the views of the various authors impels the authors themselves to consider whether such an interpretation is correct or not. By doing this, it may lead them to stand by their position, to re-interpret the position put forward, or to show how they have modified that position through the years. There is an advantage of having students try to reconstruct the positions over having the authors do it. To a great degree, the students will only react to the written materials as it might be generally interpreted by the public. The authors, on the other hand, might be inclined to read new meanings into their work saying that it expresses what they really meant. Or, authors might not see the discontinuity between their earlier writings and their later writings. When the students formulate the position, the authors may likely be surprised that it can be interpreted in such a stark way -- perhaps different from what they thought they had said or from views they currently hold.

Pitts presents an accurate description of the written views of Rippey on transactional evaluation. She offers some hypotheses about how those written conceptions might have been affected or might have been modified or reinforced based upon a reading of The New Rhetoric (TNR). In a number of places, she presents the specific issues of transactional evaluation. She hypothesizes

that there are a number of "parallels that exist between the transactional evaluation process and the process of argumentation as it is described in TNR."

Apparently, Rippey finds very little to disagree with in the interpretation made by Pitts, but an examination of his paper reveals that he is much more interested in the application of TNR to his current research and current activities. Rippey maintains that transactional evaluation is really a form of argumentation. Thus, he considers TNR as providing a valuable tool for developing various techniques and methodologies in transactional evaluation. For example, he says that he is encouraged by Perelman's thesis on argumentation to further develop his theory of confidence testing. Here Rippey illustrates the use of TNR as a tool in theory development itself.

A related point, in his example of a writing project, Rippey demonstrates that he views transactional evaluation as a form of argumentation influencing behavior. What he has done is to become more clear about the nature of his activity as a form of argumentation. In essence, he has become more conscious and more aware of the fact that his transactional evaluation activities were a form of argumentation. In becoming aware of that, he changed the things that he did and, the way in which he did them.

Rippey was also influenced in a "profound way" by the necessity of establishing a "community of minds." Apparently, Rippey took this to mean establishing the universal audience (in Perelman and Olbrechts-Tyteca's sense). Unfortunately, Rippey offers us no reasons as to why he feels that the universal audience is such an important concept of ad hominem which would not have required Rippey to establish the universal audience sense of a community of

minds. In other words, if one thought of transactional evaluation as argument ad hominem all that the evaluator would have to do would be to address ~~the group of individuals he was working with. He would not have to refer~~ their beliefs and values to any kind of universal audience at all. From this interchange we can conclude that Rippey must, in some sense, value a view of transactional evaluation which appeals to the universal audience rather than a transactional evaluation which appeals to a particular audience. This discussion is a part of the comments made by Rippey related to the use of project monitoring committees and the determination of what individuals should not be a part of such a committee. In essence, by this definition and his recourse to the universal audience, a further elaboration of Rippey's views of transactional evaluation has occurred. Up to this point, Rippey, perhaps, was unclear about whether he should address the universal audience when he does transactional evaluation or whether he should address ad hominem the audience.

An insight into the personal sequence of evaluation theory development used by Rippey is provided in his paper. Rippey believes that TNR can provide criteria for deciding who should be on the monitoring committees. Apparently Rippey believed prior to reading TNR that any protagonist or antagonist should be allowed to be on the committee. Now he feels, however, that a stronger condition must be satisfied in order for a person to be on one of these monitoring committees -- basically he must join or be a member of the community of minds. In other words, Rippey believes that members of monitoring committees must be ready to see things from the viewpoint of the interlocutor and restrict themselves to what he admits and the given effect of ones' own beliefs. Rippey's

reading of TNR has led him to see an aspect of his own work in a new way. In general, it is helpful to review work on problems that overlap or are similar to one's area of inquiry. In Rippey's case he came up with a promising lead or hypothesis. Rippey regards this view as a hypothesis or as a conjecture to be tested out or to be extended by further work. Regarding a fruitful line to explore he notes, "I am anxious to explore the effects of the establishment of such a community of minds...perhaps there is no point in preceding before this has been done." The general point about evaluation theory development that we are drawing from this example is that while reading books in related fields one often comes upon an idea which becomes a hypothesis in his own field of inquiry. This promising lead deserves further elaboration and the theorist prepares to take steps to test out and gather evidence for or against its inclusion in the theorists previously established conceptual framework. Meanwhile, the lead is considered in the tentative and speculative category.

A striking example of the multi-faceted interest of individuals and the extent to which readers may form perceptions of the views of theorists, based only upon one or a limited number of the faceted views of theorists, is presented in the Rippey paper. In his discussion of "confidence testing" Rippey illustrates another feature of theory development namely that a given theorist can be interested in the global or general principles of a theory and at the same time be interested in more particular specifics. Although Rippey is interested in developing transactional evaluation as a theory he also spent a considerable amount of time developing and using a particular methodological technique called confidence testing. The important conclusion

that we draw from this is that theoreticians in general are concerned at different levels with different aspects of the evaluation process. There is no reason to expect that all theoreticians will be equally interested in all the different levels, or to expect that all theoreticians will even concern themselves with different levels.

Guba's paper also provides promising insights into the nature of evaluation theorizing. One point made by Guba relates to what might be called the control of theorizing. Guba admits that it is always difficult for the developer of a theory to know at any given stage the following: (1) what his current views are; (2) what assumptions he is making; (3) what data is arbitrarily included or excluded; and (4) which opinions he takes into account and which he eschews. This is a very important point about an important feature of theorizing. That is, to some degree we are only loosely aware of various features about our thinking. It would have been helpful here if Guba had given us some recommendations of procedures for organizing and safeguarding our thinking. The general point cannot be too overstated that our thinking in the early stages of development is very informal, very tentative, and in no sense represents a kind of a deductive system. Perhaps if Guba were given more time, he would have presented some of the principles that should be followed so that one periodically reviews one's own work. In general, a theoretician must make a determined and persistent effort in order to put his ideas into a systematic and up-to-date whole. A related point is presented by the firsthand confrontation of a theorist - (Guba) with his earlier theoretical statements and the extension and examination of them by another person (Lindheim). Since it is always difficult for

the developer of a theory to know where his mind is, occasionally the prominent presentation of one's own views played back is sufficient influence to provide a basis for disavowal of previous points of view. It may even provide an occasion for reemphasis or restructuring by the theorist in an attempt to clear up any kinds of problems he had in communicating his ideas. Both of these things could be represented in Guba's reaction to Lindheim. The theoretician should also use the more informal occasions in which his colleagues and students give their reactions to his work. While such practice may be helpful in creating good working relationships, it is most helpful in the construction of theories.

An allied comment is the importance of recognizing the difference between a person and his theory. Theories, in essence, have an existence of their own. Guba responded to Lindheim's paper by emphasizing how much his views had changed. Apparently, Guba thought that it was his responsibility to clarify where he is today i.e., what his views about evaluation are today. On the other hand, Lindheim's clarification of the older views of Guba could be very valuable for theory development even though, or in spite of the fact that, Guba now recants or disavows them. There are at least four ways in which such work could be useful.

First, Lindheim's clarification of Guba's earlier views acts as a good to the theorist in examining his own current views and prompts him to make clear the position that he now holds. Such a clarification often allows the public and the writer himself to determine whether a significant charge was occurred on the theorizing over time.

There is the second way in which such clarification is useful. Since

theories are constructed in reaction to or in the context of previous theorizing, a complete understanding of a particular theory must depend in part on understanding the historical development of theories. It is not an accident that explanations of viable theories often take the form of showing their historical developments. House points out that Guba's writings and views have had considerable impact on theorizing over the past ten years. Therefore, Lindheim's work would help us to understand in perhaps a deeper way the nature of that work which had such a great historical influence.

A reconstruction of a theoretical position -- such as the one Lindheim gave of Guba's view -- offers a third use for theorizing. A reconstruction of a theoretical position such as Guba's may well provide clarification and explication for those who still stand by Guba's position. Lindheim's reconstruction could lead them into refining and developing their position. A reconstruction of a theoretical position also may enable those who hold differing positions to come to see something new or something which needs more emphasis in their own position.

Finally, Lindheim's work on Guba's older view might present a new interpretation of the view which might win acceptance today. This new interpretation of the position might well withstand criticisms of the older interpretation. Thus there are good reasons to pay due attention to the reconstruction of theoretical positions, even positions which are no longer held to be plausible by the author or by the majority of theoreticians. Such reconstruction can lead to fruitful developments in one's own theorizing.

An interesting sidelight and added perspective into Guba's former views of evaluation are provided by his current disavowal of that former theoretic

position. Here the principle is that one should attend to disavowals because they can be illuminating. Disavowals are especially illuminating when the theoretician presents his reasons for disavowals. In the disavowal, Guba implies that certain functions were fundamental to his previous points of view. By the nature of the assumptions that are cited, he clarifies and lets the reader have further insight into the conception of evaluation he held. His emphasizing the assumptions in such a disavowal will allow us to place the assumptions in their proper perspective in the original theory. Although the theorist originally formulated the view, he may never have been explicit about the particular assumptions and the logical role of the theory. By now emphasizing the assumptions (which he indicates are no longer tenable) he suggests that they played a major or a fundamental role in the nature of the previous conception.

The Lindheim-Guba interchange illustrates a feature similar to the first exchange of ideas. Namely, when House's writings convince Guba of the importance of the fact that there are various audiences to an evaluation each of which is persuadable by a different kind of evaluation argument, Guba then tries to fit it into his theoretical position. Apparently this fact is incompatible with other aspects of Guba's position for he goes on to draw various hypotheticals or conjectures from this fact. For example, he says that the notion of audiences seems to imply the task of the evaluator is to construct a rhetoric that is appropriate to the variety of audiences which exist. Having accepted the notion of audience and of shaping a rhetoric for the audiences, Guba immediately becomes concerned about the issue of what constitutes ethical behavior in that context. Does ethical behavior

include slanting data for different audiences, picking out what they want to hear and telling them that? The general feature involved here is that ~~accepting one fact or coming to appreciate the different forms of one fact~~ will often lead a theoretician to raise questions which he previously hadn't even suspected existed. Therefore, persuasion on a certain point ends up in furthering theory development by promoting leading questions. Such a dialogue as the one given by Guba presents an important picture of theorizing -- it is rather informal and dynamic in its nature.

The Guba-House interchange also illustrates another feature of theorizing, for we can assume that Egon Guba had not given serious consideration or had supposed that the ethical questions were resolved until he really appreciated the force of House's remarks on the issue of rhetoric for various audiences. These remarks and agreement with this view forced him to rethink his stand on the ethical questions. We can therefore assume that previously Guba had supposed that all of his evaluation reports were aimed at the universal audience -- in other words, that his evaluation reports were on the side of truth and justice and objectivity. Having accepted House's point that most evaluation reports are directed toward particular audiences -- subject to the use of rhetoric for argumentation -- Guba was led to re-evaluate certain of his beliefs. In particular, he re-evaluated the ethical questions related to argumentation directed toward particular audiences.

Curiously enough Wolf also comes to dwell on the ethical considerations. Wolf's reply to TNR represents another interesting feature of theory development. Wolf perceives certain elements in TNR, as presented in House's paper or in Daillak's, as offering the possibility of corrupting or corroding his

own theory of evaluation. He then went on, provoked by these comments, to give an elaborate defense or at least a punitive defense of the implications of accepting one of the elements of TNR into his own point of view. By reacting to a particular point in TNR, Wolf indirectly illuminates and disassociates his own theoretical point of view from that element of TNR. He, in essence, helps to clarify one facet of his own theory. He shows for example, by focusing on partisan advocacy in contrast to his own judicial model--designed to secure impartial clarification--that this aspect of Perelman's thesis contrasts quite sharply with his own view of evaluation. He then goes on to provide a deeper understanding of his own view by claiming that the reason that impartial clarification is so valuable or so fundamental is that it will lead to informed decisions and significant changes. By elaborating this point in his theory, Wolf really presents one of the fundamental premises in his whole theory--namely, that the ultimate goal of evaluation is to reach informed decisions which lead to significant changes.

Wolf's reply to Popham's remarks on adversary evaluation as a possible criticism of his own position illustrates other strategies which are open to the theoretician. Instead of disavowing his position because of the Popham criticism, Wolf draws the distinction between adversary and JEM evaluation. This allows him to admit the force of Popham's criticism while he can stand by his own JEM position. Having been lead to address Popham's criticism, Wolf goes on to emphasize in a new way, his view of the central aspect of JEM. As he says, at the heart of JEM is the view that the more points of view there are on the matter the more likely it will be that the truth is found. Thus, consideration of various viewpoints is valued because

it is a means of finding the truth. Wolf adds that JEM provides conditions which make argumentation on the facts possible. Apparently he feels JEM's conditions rule out, or rule out as unlikely, the process of adversary evaluation.

The examination of Wolf's reply to Dailak presents insight into some of the less understood features of Wolf's JEM theory. For example, it appears Wolf believes that his process of JEM amounts to various members of the committee judging factual matters. When he talks about impartial evaluation he also talks about cross-checking and cross-validation and trying to confirm hypotheses. This leads one to conclude that in a way analagous to the law procedure, Wolf believes that evaluation (his evaluation procedure) is basically aiming at determining statements of fact by using various groups who view those facts from different perspectives. So, he will be optimistic about resolution because he believes that the use of JEM will allow people to agree on empirical facts. In fact, Wolf specifically disavows argumentation and persuasion as essential elements of his theoretical point of view and provides added insights into his conception of JEM by drawing the distinction between JEM and adversary evaluation techniques. His use of the word "clarification" is to describe one of the functions of the JEM process, and his language such as "they view the world from a different frame of reference" leads one to believe that Wolf holds that the outcome of JEM will be agreement on matters of fact. In his view, adversary techniques really amount to arguing for a favor or pleading for special points of view. Somehow, Wolf believes his JEM technique is very distinct from adversary evaluation in that it enhances resolution of the empirical matters

of fact. Thus we see quite vividly that by presenting Wolf with the opportunity to disavow elements of TNR as inappropriate and inapplicable to his evaluation theoretic point of view, we gain greater understanding into the nature of his own formulation.

In this essay, we have been attempting to draw out various principles which govern evaluation theorizing. The focus which has been used is an examination of the replies and reactions of several theoreticians to a reconstruction or interpretation of their theoretical positions by the "role players." We have not nor do we mean to say that our interpretation of their theorizing is final. What we have tried to show is the plausibility of certain principles underlying their theorizing. As papers by the role players were a stimulus for re-affirmation, re-conception, or disavowal of one's previous theoretical formulation, so do we expect that this analysis will further serve as a factor affecting the dynamics of evaluation theorizing. As we have tried to show, theorizing is an informal and dynamic process.

References

- Braybrooke, D., & Lindblom, C. A strategy of decision. New York: Free Press, 1963.
- Glass, G. V. Educational product evaluation: A proto-type format applied. Educational Research, 1972, 1(1), 7-10, 16. (Reprinted in House, 1977.)
- Guba, E. G. The failure of educational evaluation. Educational Technology, 1969, 9(5), 29-38.
- Guba, E. G. Problems in utilizing the results of evaluation. Journal of Research and Development in Education, 1975, 8(3), 42-54.
- Guba, E. G., & Horvat, J. J. Evaluation during development. Viewpoints. (The Bulletin of the School of Education, Indiana University), 1970, 46(2), 21-45.
- House, E. R. The conscience of educational evaluation. Teachers College Record, 1972, 73(3), 405-414.
- House, E. R. The logic of evaluative argument. CSE Monograph Series in Evaluation, No. 7. Los Angeles: Center for the Study of Evaluation, University of California, 1977.
- Perelman, Ch., & Olbrechts-Tyteca, L. The new rhetoric: A treatise on argumentation. (J. Wilkinson & P. Weaver, trans.). Notre Dame, Ind.: University of Notre Dame Press, 1969. (Originally published, 1958.)
- Popham, W. J., & Carlson, D. Deep dark deficits of the adversary evaluation model. Educational Researcher, 1977, 6(6), 3-6.
- Rippey, R. M. Transactional evaluation and the improvement of instruction. Educational Technology, 1977, 17(2), 7-11.
- Rippey, R. M. Studies in transactional evaluation. Berkeley, California: McCutchan, 1973.
- Scriven, M. Educational product re-evaluation. Educational Researcher, 1972, 1(5), ii, 2, 12. (Reprinted in House, 1977.)
- Wergin, J. F. The evaluation of organizational policy making: A political model. Review of Educational Research, 46(2), 1976, 75-115.
- Wolf, R. L. The application of judicial concepts to educational evaluation. Unpublished doctoral dissertation, University of Illinois, 1973.
- Wolf, R. L. Evidence: Educational evaluation and the metaphors of law. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D. C., April 2, 1975.

Wolf, R. L., & Tymitz, B. L. Enhancing policy formulation through naturalistic/judicial inquiry procedures: A study of the individual education program component of Public Law 94-142. Unpublished manuscript, Indiana Center for Evaluation, Indiana University, 1977(a).

Wolf, R. L., & Tymitz, B. L. Towards more natural inquiry in education. CEDR Quarterly, October, 1977(b).