An Exploratory Study of Sentient Evaluators: Communication, Ethics, and Relationships in the Evaluation Process

Mary Elizabeth Piontek
Western Michigan University

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Educational Assessment, Evaluation, and Research Commons, and the Educational Leadership Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/1861
AN EXPLORATORY STUDY OF SENTIENT EVALUATORS:
COMMUNICATION, ETHICS, AND RELATIONSHIPS
IN THE EVALUATION PROCESS

by

Mary Elizabeth Piontek

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Doctor of Philosophy
Department of Educational Leadership

Western Michigan University
Kalamazoo, Michigan
August 1994
AN EXPLORATORY STUDY OF SENTIENT EVALUATORS: COMMUNICATION, ETHICS, AND RELATIONSHIPS IN THE EVALUATION PROCESS

Mary Elizabeth Piontek, Ph.D.
Western Michigan University, 1994

The purpose of this exploratory study was to develop hypotheses about the ways evaluators involve stakeholders, clients, and audiences in the evaluation process and communicate their findings to these groups. A sentient evaluator was defined as an evaluator who takes an active role in promoting the program that s/he is investigating and deliberately seeks to represent underrepresented stakeholders. This type of evaluator has her/his own ethical or moral agenda through which s/he views the program being evaluated and takes a formative perspective on program development and improvement.

Evaluators (n=12) believed to be sentient because they displayed certain characteristics were compared to randomly selected members of the American Evaluation Association (n=7) on two dimensions: (1) the ethical dilemmas that they encounter in the evaluation process, and (2) how they involve and communicate to various stakeholders/clients/audiences.

This study developed detailed descriptions of the
sentient role in terms of communicating and reporting efforts, the involvement of various groups in the evaluation process, and ethical considerations, in order to better understand its implications for the profession of evaluation. The following hypotheses were developed for future research based on the case interviews conducted in this study:

Hypothesis 1: Sentient evaluators differ from other evaluators in terms of the context they ascribe to evaluation (macro vs. micro-level of focus); Hypothesis 2: Sentient evaluators differ from other evaluators in terms of evaluation's role in promoting social criticism; Hypothesis 3: Sentient evaluators differ from other evaluators in terms of the involvement of stakeholders in the evaluation process; Hypothesis 4: Sentient evaluators differ from other evaluators in terms of the length, breadth, and depth of involvement in the project/organization being evaluated; and Hypothesis 5: Sentient evaluators differ from other evaluators in terms of evaluation's need to have theoretical discussion of values (how they affect the evaluation process and how they are communicated to stakeholders/clients/audiences).
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
An exploratory study of sentient evaluators: Communication, ethics, and relationships in the evaluation process

Piontek, Mary Elizabeth, Ph.D.
Western Michigan University, 1994

Copyright ©1994 by Piontek, Mary Elizabeth. All rights reserved.
# TABLE OF CONTENTS

## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>v</td>
</tr>
</tbody>
</table>

## CHAPTER

I. **OVERVIEW**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>4</td>
</tr>
<tr>
<td>Conceptual Framework</td>
<td>6</td>
</tr>
<tr>
<td>Limitations</td>
<td>9</td>
</tr>
</tbody>
</table>

II. **REVIEW OF LITERATURE**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Evaluation Models</td>
<td>12</td>
</tr>
<tr>
<td>First Wave</td>
<td>12</td>
</tr>
<tr>
<td>Second Wave</td>
<td>13</td>
</tr>
<tr>
<td>Evaluation Roles</td>
<td>14</td>
</tr>
<tr>
<td>Summative and Formative</td>
<td>15</td>
</tr>
<tr>
<td>Internal/External</td>
<td>15</td>
</tr>
<tr>
<td>Evaluator Roles</td>
<td>16</td>
</tr>
<tr>
<td>Expert</td>
<td>17</td>
</tr>
<tr>
<td>Power Merchant</td>
<td>18</td>
</tr>
<tr>
<td>Interpreter</td>
<td>18</td>
</tr>
<tr>
<td>Mediator or &quot;Broker&quot;</td>
<td>19</td>
</tr>
<tr>
<td>Facilitator</td>
<td>20</td>
</tr>
<tr>
<td>Consultant/Educator</td>
<td>20</td>
</tr>
<tr>
<td>Personal Coach</td>
<td>21</td>
</tr>
<tr>
<td>Investigator</td>
<td>21</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Theorist/Emancipator</td>
<td>22</td>
</tr>
<tr>
<td>Connoisseur/Critic</td>
<td>22</td>
</tr>
<tr>
<td>Social Critic/Critical Ethnographer</td>
<td>23</td>
</tr>
<tr>
<td>Sentient</td>
<td>24</td>
</tr>
<tr>
<td>Ethical Considerations in Evaluation</td>
<td>26</td>
</tr>
<tr>
<td>Communicating and Reporting Evaluation Findings</td>
<td>31</td>
</tr>
<tr>
<td>Involvement in Utilization of Findings</td>
<td>33</td>
</tr>
<tr>
<td>Political Context</td>
<td>33</td>
</tr>
<tr>
<td>Evaluator Credibility and Relationships</td>
<td>34</td>
</tr>
<tr>
<td>III. METHODOLOGY AND DATA ANALYSIS</td>
<td>36</td>
</tr>
<tr>
<td>Participants</td>
<td>36</td>
</tr>
<tr>
<td>Choice of Methodology</td>
<td>40</td>
</tr>
<tr>
<td>Development of Interview Protocol</td>
<td>42</td>
</tr>
<tr>
<td>Interview Process</td>
<td>44</td>
</tr>
<tr>
<td>Data Analysis Overview</td>
<td>47</td>
</tr>
<tr>
<td>Transcription</td>
<td>47</td>
</tr>
<tr>
<td>Coding Scheme</td>
<td>49</td>
</tr>
<tr>
<td>Analysis of Coded Data</td>
<td>54</td>
</tr>
<tr>
<td>Credibility, Transferability, Dependability, and Confirmability of the Study</td>
<td>56</td>
</tr>
<tr>
<td>IV. PRESENTATION OF DATA AND FINDINGS</td>
<td>61</td>
</tr>
</tbody>
</table>
Table of Contents—Continued

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview</td>
<td>..................................................</td>
<td>61</td>
</tr>
<tr>
<td>Demographic Information</td>
<td>...........................................</td>
<td>62</td>
</tr>
<tr>
<td>The Purpose of Evaluation</td>
<td>........................................</td>
<td>64</td>
</tr>
<tr>
<td>The Model/Method of Inquiry Used</td>
<td>....................................</td>
<td>68</td>
</tr>
<tr>
<td>The Role of Stakeholders/Clients/Audiences in the Evaluation Process</td>
<td>.................</td>
<td>77</td>
</tr>
<tr>
<td>The Communicating and Reporting of Evaluation Findings</td>
<td>....................</td>
<td>84</td>
</tr>
<tr>
<td>Ethics and Social Responsibility</td>
<td>..............................</td>
<td>91</td>
</tr>
<tr>
<td>Future Issues for Evaluation</td>
<td>............................</td>
<td>102</td>
</tr>
<tr>
<td>V. CONCLUSIONS</td>
<td>...........................................</td>
<td>105</td>
</tr>
<tr>
<td>Implications</td>
<td>........................................</td>
<td>105</td>
</tr>
<tr>
<td>Sentient</td>
<td>............................................</td>
<td>113</td>
</tr>
</tbody>
</table>

APPENDICES

<table>
<thead>
<tr>
<th>APPENDICES</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Interview Protocol</td>
<td>........................................</td>
<td>117</td>
</tr>
<tr>
<td>B. Cover Letter and Overview</td>
<td>..................................</td>
<td>120</td>
</tr>
<tr>
<td>C. Informed Consent Form</td>
<td>.......................................</td>
<td>126</td>
</tr>
<tr>
<td>D. Timetable with Notes</td>
<td>.........................................</td>
<td>130</td>
</tr>
<tr>
<td>E. Site Form</td>
<td>........................................</td>
<td>139</td>
</tr>
<tr>
<td>F. Member Checking Form</td>
<td>.........................................</td>
<td>141</td>
</tr>
<tr>
<td>G. Matrix of Evaluation Roles</td>
<td>......................................</td>
<td>143</td>
</tr>
<tr>
<td>H. Coding Book</td>
<td>........................................</td>
<td>145</td>
</tr>
<tr>
<td>I. HSIRB Protocol Approval</td>
<td>.......................................</td>
<td>150</td>
</tr>
</tbody>
</table>

BIBLIOGRAPHY | ........................................ | 159 |

iv
LIST OF FIGURES

1. Data Analysis Process .......................... 48
2. Coding Scheme - First Round ..................... 51
3. Coding Scheme - Second Round .................... 52
4. Purpose of Evaluation ............................ 67
5. Persons Involved in Preparing Evaluation Questions .................................................. 82
6. Point of Emphasis of Stakeholder Involvement .. 83
7. Social Responsibility of Evaluation in Terms of Reeducating and Discussing/Introducing Notions of Participation ......................... 102
8. Need to Have Theoretical/Grounded Pragmatic Dialogue or Discussion of Values .................. 102
9. Resistance to Changing Evaluation Findings/Validation .................................................. 102
10. Sensitivity to Using Personal Information ..... 102
CHAPTER I

OVERVIEW

If one looks at the evaluation\(^1\) literature in terms of the many different roles that an evaluator can play and the different ways evaluation can be conducted, including its purpose/focus, activities/methodologies, and relationships with clients/stakeholders/audiences, one wonders how these numerous choices/alternatives play out in the actual behavior of evaluators. Little has been written about the worldview/framework of evaluators who take an active role in counselling the program that they are investigating, deliberately seeking to represent underrepresented stakeholders or audiences. This type of evaluator may have her/his own ethical agenda through which s/he views the program being evaluated and tends to have a formative\(^2\) perspective on program development and improvement. The researcher has defined this evalua

\(^1\) Evaluation is defined as the "systematic investigation of the worth or merit of an object." (Joint Committee, 1994). The object can be a program, project, curriculum or instructional material, or product, etc. An evaluator is someone who conducts, directs, or engages in an activities related to an evaluation, formal or informal.

\(^2\) Refer to Chapter II, Evaluation Roles, for discussion of summative and formative evaluation.
tor role as "sentient."³

While factors about reporting/communicating evaluation findings have been addressed in the literature, there has been no systematic investigation into whether identifiable types of evaluators may exist that would influence how the evaluator relates to program staff, especially in terms of communicating with and involving them in the evaluation process.

This exploratory study draws from this body of knowledge on communicating/reporting, as well as three of the more salient issues of discussion in evaluation today: collaboration as a means of involving audiences, clients, and stakeholders to increase the relevance and utilization of evaluation processes and findings; alternative means of communicating and reporting evaluation findings; and the development of ethical standards for

³The term "sentient" is defined in Webster's College Dictionary as "having the power of perception by the senses, conscious; characterized by sensation and consciousness" (1991, p.1222). I have chosen this terminology to define evaluators who possess certain characteristics related to how they conduct evaluations, who they involve in the evaluation process and how they involve them, and their views of the purpose(s) of evaluation. A fuller definition of the sentient evaluator is presented in the Conceptual Framework section, Chapter 1. This is not a term that these evaluators use to define or label themselves. I have specifically chosen to create a new term or construct in this study in order to capture the differences between the characteristics that I have identified as "sentient" and the other evaluator roles that have been presented in the literature.
the profession of evaluation.

Evaluators face circumstances which continue to speak to the need for a better understanding of how evaluation practice might be made more efficient and effective. For example, evaluators (and those being evaluated) are experiencing (a) increased demands from state and federal governments, and the public for accountability and effectiveness in the domains of education and business; (b) increased involvement of parents and other audiences in the accountability/evaluation process; and (c) increased acceptance of non-traditional evaluation methods (i.e., case study, ethnographic, qualitative designs) which require different, innovative reporting styles (Torres, Preskill, & Piontek, 1992a; 1992b).

This exploratory study focuses on examples (b) and (c), with special attention to how audiences, clients, and stakeholders are involved in the evaluation process.

This exploratory study builds from the findings of a Spring, 1992 study of a random sample of 343 members of the American Evaluation Association (Torres, Preskill, & Piontek, 1992a, 1992b) in which evaluators described aspects of format, content, and process which contributed to successful communicating and reporting: (a) formats providing for easy/quick assimilation of information; (b)
content addressing both positive and negative findings, qualitative and quantitative aspects of the program, and specific recommendations for change; and (c) processes involving clients/audiences/stakeholders in the evaluation from its beginning.

These respondent evaluators also identified what followed when they used the strategies listed above. The most compelling of which was collaborating with clients/audiences/stakeholders in the conduct of the evaluation. Involving stakeholders in the evaluation process, they reported, increases the credibility of the effort, enhances their understanding of both the evaluation and the program, and increases the possibility that action will be taken on the basis of the findings (Torres, Preskill, & Piontek, 1992a; 1992b).

Purpose

The overall purpose of this study is to look at the ways evaluators involve stakeholders, clients, and audiences in the evaluation process and communicate their

---

4 Stakeholders are those persons who may be directly affected by the evaluation results - they have a 'stake' in the program or its evaluation; a client is the specific agency or individual who requests the evaluation; audiences are individuals, groups, and agencies who have an interest in the evaluation and receive evaluation results (Joint Committee, 1994; Worthen & Sanders, 1987, pp. 165-6).
findings to these various groups. It focuses specifically on evaluators, labelled by this researcher as "sentient," the programs and/or stakeholders, clients, and audiences that they evaluate, and the ethical dilemmas that this worldview/framework poses in (a) how they involve their stakeholders, clients, and audiences in the evaluation process (context or setting of the evaluation activities) and (b) how they communicate, especially in terms of the language they use and the "voices" or perspectives of various groups that they attempt to capture, including their use of format (physical organization of documents/type of medium used) and content (topical areas).

Evaluation approaches based on anthropological and feminist theories about research, its purpose, and its "authors" or "voices" are just beginning to emerge. Essentially a sentient evaluator, in this researcher's definition, takes an active role in counselling the program that s/he is investigating, deliberately seeking to represent underrepresented stakeholders or audiences. This type of evaluator may have her/his own ethical

5 The term "sentient" is used because it best captures the nuances of the evaluation framework/viewpoint/philosophy discussed here. No new terminology emerged from the interview data that these evaluators themselves use to define who they are and what they do.
agenda through which s/he views the program being evaluated and tends to have a formative perspective on program development and improvement.

The intent of this study is neither to promote or criticize this worldview/framework, but rather to develop detailed descriptions about it in terms of communicating and reporting efforts, the involvement of various groups in the evaluation process, and ethical considerations, in order to better understand its implications for the profession of evaluation. In addition, preliminary hypotheses will be generated that may form the basis of future research in this area.

Conceptual Framework

Sentient evaluators, as defined here, have the following characteristics which separate them from other groups of evaluators. It is essentially their worldview, philosophy, or framework under which they conduct themselves and their professional evaluation practice that separates them. Although each sentient evaluator may not embody all of the characteristics listed below, in some sense their overall focus and belief system reflect these. Essentially, in this study, the term sentient is used as a new construct, emerging out of the synthesis of other roles of an evaluator and the overall framework,
under which they engage in evaluation activities.

Essentially a sentient evaluator takes an active role in counselling the program that s/he is investigating, deliberately seeking to represent underrepresented stakeholders or audiences. This type of evaluator may have her/his own ethical agenda through which s/he views the program being evaluated and tends to have a formative perspective on program development and improvement.

These evaluators to some extent assume that the world is complex and that programs are embedded in social and political processes and organizations. Participants are involved with the sentient evaluator in designing, conducting, and disseminating the evaluation.

This worldview/framework/philosophy synthesizes and goes beyond nine recently discussed roles in evaluation: (1) the interpreter's revealing of multiple perspectives and realities (Mishler, 1979; Noblit & Eaker, 1987; Spicer, 1976); (2) the mediator-broker's movement between social groups and structure (Greene, 1992; Noblit & Eaker, 1987; Torres, 1991); (3) the facilitator's involvement of participants and marginalized groups (Greene, 1992); (4) the consultant-educator's educative-reflective opportunities (Cousins & Earl, 1992; Torres, 1991); (5) the coach's drive to increase effectiveness/efficiency (Hendricks, 1993); (6) the investigator's
search for hidden meanings (Smith, 1992); (7) the emanci-
pator's discussion of ideological content (Greene, 1988; 
Noblit & Eaker, 1987); (8) the connoisseur's appraisal of 
the evaluative situation (Eisner, 1979); and (9) the 
social critic-critical ethnographer's challenge of ethi-
cal commitments (Greene, 1992; Thomas, 1993). Although 
the sentient is similar to these other roles, the synthe-
sis results in a slightly different overall framework.

While challenging the social, political, personal 
structures and beliefs in the evaluation context, the 
evaluator, identified as sentient, also simultaneously 
challenges her/his own structures and beliefs. S/he 
simultaneously reveals her/his own agendas, biases, 
beliefs, and understandings as s/he reveals those of the 
context. Closely aligned to cultural feminism, which 
emphasizes a holistic and collective orientation to world 
and work experiences, s/he creates a kind of praxis, with 
continuous cycles of critique, knowledge construction, 
and social action.

The sentient evaluator may be reflected in some of 
the concerns behind creating "morally engaged" evalua-
tion. The sentient evaluator might agree that human 
inquiry cannot be divorced from historical, moral, and 
political locations and that politics and science are 
intertwined in evaluation practice. The sentient evalua-

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
tor may also agree that those who conduct research within an evaluative context are obligated to consider the moral implications that their findings may carry for those who are affected by the program. These evaluators acknowledge to their audiences their position on issues of social justice by engaging in self-conscious reflection of their values and articulating these values.

Sentient evaluators are self-critical, identify the moral issues and work to resolve moral problems in their own work, and behave in morally responsible ways. They might say that being ethical in program evaluation is a process of growth in understanding, perception, and creative problem-solving ability that respects the interests of individuals and society. These evaluators help the respondents cope with new knowledge, creating situations without deception in which the evaluator and evaluated are involved in the construction of reality.

Limitations

This study is exploratory in that it is not attempting to test any predetermined hypotheses, but rather is creating a preliminary description of sentient evaluators and how they differ from other evaluators in the profession. From this preliminary description, tentative hypotheses will be developed that may direct future re-
search.

The conclusions of this study are based on self reports of a small number of cases chosen because they met the selection criteria. There is no basis in this study for generalizing beyond the cases who were studied.

The context of this study is largely affected by the annual Meeting of the American Educational Association. The title of the 1993 annual meeting in Dallas, Texas was "Empowerment Evaluation" and many of the presentations specifically targeted this topical area. Thus, those persons who attended were probably interested in discussing this topic (and perhaps had prepared presentations/papers on this topic for the meeting). Although the persons interviewed in this study were not limited to only those persons who attended this meeting, fourteen interviews were conducted in Dallas during the four day meeting, and the responses of those persons interviewed may have been affected by the tone of the meeting, the issues being presented/discussed, and even the preliminary discussion about the focus of the 1994 annual meeting, "Social Justice Evaluation." Thus, had these interviews been conducted at another time, in another setting, the responses might have been different.

One main weakness of this study is that it is based on only the "word" of the participants in both groups as
to whether they actually "practice what they preach," because the researcher did not interview any of the clients/stakeholders/audiences that they have worked with, nor were existing evaluation reports, videos, memos, minutes of meetings, or other types of documentation of evaluation activities and communicating/reporting strategies collected. The collection of such documents and the contacting/interviewing of the stakeholders would be a complex and exhausting task, both legally and logistically, certainly not impossible, but considered by the researcher to be beyond the scope of this study.
CHAPTER II

REVIEW OF LITERATURE

Evaluation Models

First Wave

Evaluation is a field that developed in the last 30+ years, with many different models of evaluation and evaluator roles evolving within it. Essentially, each model grew from the evaluative needs and focus of its time, as well as the dominant research paradigm being applied to social research—usually discussed in terms of positivism or post-positivism (Herman, Morris, & Fitz-Gibbon, 1987).

Models of the late 1960s and early 1970s addressed the formative needs of large-scale curriculum developers, the summative needs of sponsors/funders, and the accountability needs of federal policy-makers. These models utilize experimental methods, standardized data collection, and large samples, with the belief that scientific measurement produces hard, unequivocal evidence of program failure or success, essentially, evaluation as positivistic research (Herman, Morris, & Fitz-Gibbon, 1987). Focusing on cause-effect relationships, based on
the programs' goals and issues, these models provide
decision-makers and policy-makers with standardized,
aggregated information about pre-determined program
processes and outcomes (Herman, Morris, & Fitz-Gibbon,
1987; Worthen & Sanders, 1987). Two of the more prominent
models within this group are (1) Goal-oriented evaluation
(Bloom, Hastings, & Madaus, 1971; Popham, 1975; Tyler &
Smith, 1942) and (2) Evaluation Research (Campbell &
Stanley, 1966; Campbell, 1969; Rossi & Freeman, 1985).

Second Wave

Some believe that these first models neglect the
evaluation context and are insensitive to variations in
local programs. Thus a new group of models emerged that
emphasized "responsiveness" to local settings and partic-
ipants/stakeholders (Herman, Morris, & Fitz-Gibbon, 1987;
Worthen & Sanders, 1987). This second group of models
assume that the world is complex and that programs are
embedded in social and political processes and organiza-
tions. Participants are involved in designing, conduct-
ing, and disseminating the evaluation. Appropriately,
these models utilize naturalistic, qualitative methods in
order to provide an indepth understanding of unique
program characteristics, (essentially post-positivistic
research). Models which fit into this group include (a)
Responsive-Participatory Evaluation (Cousins & Earl, 1992; Stake, 1975); (b) Naturalistic Inquiry (Guba, 1969; Guba & Lincoln, 1981, 1989); and (c) Democratic Evaluation (MacDonald, 1974, 1976). However, these models lack credible ways of generalizing across sites and thus hold clear problems for accountability or decision-making needs.

In conjunction with the development of responsive models, evaluation began to attend to the utility of evaluation findings. The impact of evaluations and evaluation findings was questioned, and models stressing the importance of evaluation's contribution to social policy, planning, and practice evolved. These include (a) Utilization-Oriented Evaluation (Alkin, Daillak, & White, 1979; Patton, 1986); (b) Advocacy-Adversary Evaluation (Wolf, 1975; Owens, 1970); (c) Management-Oriented Evaluation (Alkin, 1969; Stufflebeam, 1971a, 1971b); and (d) Government-Decision Making Evaluation (Wholey, 1986).

Evaluation Roles

In addition to discussing the field of evaluation in terms of models, evaluation is often "sorted" into the roles it plays in the program being evaluated.
Summative and Formative

Evaluation often takes on a formative or summative role, or both. Formative evaluation is conducted during the operation of a program where the feedback loop remains inside of the program/organization to provide evaluative information useful in improving the program or product (Scriven, 1967); its focus is on program development with an audience of program personnel. Summative evaluation is conducted at the end of a program to provide potential consumers, outside of the program/organization, with judgments about the program's worth or merit to improve utilization or recognition of the program/product (Scriven, 1967); its focus is on continuation/termination or adoption of a program with a traditional audience of potential consumers, funders, and program supervisors.

Internal/External

Another way of discussing evaluation is the "setting" or relationship of the evaluation "unit" (enterprise, function, or effort) to the client or audience. Internal evaluators operate within an organization. "The internal evaluator is an employee who holds explicit, primary responsibility for the organization's self-evaluation. The key terms here are internal, employee, and explicit responsibility" (Clifford & Sherman, 1983,
External evaluators are hired as independent consultants from outside of the organization based on negotiated contracts (Mathison, 1991). Many evaluation models can cut across these roles, depending on the timing, location, and purpose of their application.

Certainly, this overview does not cover all models of evaluation [others include Goal-Free Evaluation (Scriven, 1972) and Educational Connoisseurship (Eisner, 1976; Flinders, 1986)], but instead sets the context in which various roles of evaluators emerged, including facilitator, expert, mediator, critic, investigator, consultant, patron, connoisseur, power-merchant, translator, interpreter, theorist, and coach. Some roles of evaluators can cut across methods of inquiry, models of evaluation, and summative/formative and internal/external evaluation roles, but others promote a very particular, unique evaluative purpose, perspective, and methodology. These evaluator roles are not meant to be seen as mutually exclusive; an evaluator may play many roles simultaneously or alternate during the course of an evaluation.

Evaluator Roles

To clearly understand the evaluators labelled as
sentient in this study, it is important to identify other roles that have been discussed in evaluation literature and the characteristics that have been used to define or designate these other roles.

The eleven other evaluator roles presented here capture the most commonly discussed evaluator characteristics, responsibilities, or perspectives in the literature; this list is not exhaustive. The eleven roles presented here serve as a comparison and contrast to those characteristics identified as the "sentient" role.

Expert

The role of the expert or expert-scientist can be seen within the research/evaluation approach of positivism. Positivism supports science as the superior way of knowing, with knowledge discovered through reductionist epistemology (Noblit & Eaker, 1987). In this setting, the scientist is the expert or credible agent; objectivity is the key element (Noblit & Eaker, 1987). Human/social systems are viewed in terms of a positivistic cause-effect linear model, providing clear predictions and explanations of the relationships between educational objects and events (Smith, 1983).

The expert-scientist evaluator role in this setting is legitimized by her/his scientific expertise. Interper-
sonal skills are not emphasized; the relationship between the evaluator and evaluatees is more authoritarian with a focus on objectivity. The expert-scientist in a sense is a "patron" to the sponsor/evaluatees "clients" maintaining control over the focus/procedures of the evaluation (Noblit & Eaker, 1987).

Power Merchant

The evaluator as "power-merchant" helps an organization collect data, develop constructive information, structure instruments and reports, and develop a common language, resulting in a transfer of power (power being knowledge, information, or the ability to act upon a decision) to program staff and an increase in the use of evaluation findings (Fein, Staff, & Kobylenski, 1993). The power merchant understands the social and political processes of the organization and its context and uses the opportunities of these processes to increase program success and evaluation utilization.

Interpreter

Within interpretivism, which focuses on putting the meaning of human/social situations in relevant contexts (Mishler, 1979), the evaluator takes on the role of interpreter or "revealer". Characteristics of interpreti-
vism include (a) a grounding in the experiences of the participants, and being (b) holistic, (c) historical, and (d) comparative (Spicer, 1976). The evaluator observes and interviews the participants in order to construct an ethnography or case study of multiple perspectives and multiple realities. The evaluator serves as an interpreter or "revealer" of underlying assumptions of the situation (Noblit & Eaker, 1987).

Mediator or "Broker"

In interpretivism, the evaluator may also serve in the role of a broker or mediator, moving between different social structures and groups at times participant, at times observer, "transmitting messages or facilitating social exchanges" at the end of an evaluation (Noblit & Eaker, 1987). This role of broker can also be seen in collaborative research, where the evaluator, client, audiences, and stakeholders collaborate and negotiate the research design, the roles of all participants, and the issues in the evaluation (Greene, 1992; Noblit & Eaker, 1987). The broker is essentially a translator of information or "messages" between all parties. The mediator "provides leaders opportunities to reflect critically on the organizational context and the individuals who comprise it" (Torres, 1991, p.190).
Facilitator

Another evaluator role in collaborative research is that of facilitator of a holistic understanding of the evaluative context and promoter of participant involvement in and ownership of the evaluation process and findings (Greene, 1992). The facilitator gives voice to those who may be marginalized or underrepresented and makes possible a participatory inquiry process that "enables the joint, mutual construction and ownership of meaningful, important knowledge and the capacity to act on that knowledge" (Greene, 1992, p. 5).

Consultant/Educator

As consultant/educator, the evaluator becomes an educator on "evaluation" (Cronbach, 1980). By checking her/his impressions of the evaluation context with organizational members, the evaluator provides an opportunity for members to "reflect on their own environment and clarify it for themselves" (Torres, 1991, p. 192). By facilitating this understanding of the context, evaluation utilization and organizational change may be expedited (Torres, 1991). The consultant/educator cues in on "teachable moments" (Torres, 1991, p. 194) to develop a shared language, fulfill a pedagogical role in teaching about evaluation, and motivate primary users to partici-
pate in evaluation activities (Cousins & Earl, 1992).

**Personal Coach**

Similar to the consultant role is the evaluator as "personal coach". The formative evaluator in this role tries to make a program/project efficient and effective, working closely with the client, yet remaining objective (Hendricks, 1993). A "personal coach" evaluator can help a program manager or policy maker to (a) recognize that her program matters and that how well it performs is important; (b) clarify a program's criteria for successful performance; (c) believe that her program can significantly improve; (d) acknowledge that outside help may be necessary to improve performance; (e) assess objectively her program's current performance; (f) understand what makes his program effective; (g) consider suggestions for improving her program; (h) implement those recommendations which are accepted; and (i) continue to seek further performance improvements (Hendricks, 1993, pp.51-53).

**Investigator**

As an investigator, the evaluator utilizes the four "mental abilities" of knowledge, observation, reasoning, and intuition in an "alternatively exploratory and
confirmatory, recursively emergent process of investiga-
tive inquiry" in order to "search into, to inspect, to
explore, to confirm" information in an evaluation context
(Smith, 1992, pp.10, 3). Like a journalist or patholo-
gist, the investigator-evaluator tries to discover some-
thing hidden, through various inquiry techniques (Smith,

Theorist/Emancipator

In critical theory as a mode of evaluation, the
evaluator becomes a theorist that serves as an emancipa-
tor or "enlightener" of people by addressing the ideolog-
ical content of the beliefs that structure their lives.
Discourse and dialogue are used to promote self-reflec-
tion (Noblit & Eaker, 1987). Similarly, in participatory
evaluation, enlightenment may take place through the
active engagement of stakeholders and clients in the
total evaluation process, from design to communicating
and reporting evaluation findings (Greene, 1988).

Connoisseur/Critic

Aesthetics is a qualitative approach to evaluation,
involving both connoisseurship and criticism (Eisner,
1979). Connoisseurship involves the "art of apprecia-
tion"; criticism creates an artistic description that
others participate in vicariously, interprets the descriptions by applying "theoretical ideas", and appraises the situation by "providing conditions that lead to the improvement of the educational process" (Eisner, 1979, pp.15-16). In this role of connoisseur-critic, interpersonal skills and trust are essential (Noblit & Eaker, 1987).

Social Critic/Critical Ethnographer

The role of the social critic or critical ethnographer is similar to the connoisseur role in that it results in changes in cognition; is similar to the role of interpreter in that it provides reflection that examines culture, knowledge, and action (Thomas, 1993); and is similar to the role of enlightening-emancipatory-theorist in that it describes, analyzes, and scrutinizes hidden agendas, centers of power, and assumptions in beliefs and structures for change-oriented reflection and action (Greene, 1992; Thomas, 1993). The social critic-critical ethnographer "deepens and sharpens ethical commitments by forcing us to develop and act upon value commitments in the context of political agendas...[s/he] requires that common sense assumptions be questioned" (Thomas, 1993, pp. 2-3).

To better illustrate the differences between these
eleven evaluator roles, a matrix with each evaluator role and its distinguishing characteristics (noting overlap in characteristics where appropriate) is presented in Appendix G.

**Sentient**

The evaluator defined as "sentient" in this study is an evaluator who takes an active role in counselling the program that s/he is investigating, deliberately seeking to represent underrepresented stakeholders or audiences. This type of evaluator may have her/his own ethical moral agenda through which s/he views the program being evaluated and tends to have a formative perspective on program development and improvement.

This worldview/framework/philosophy synthesizes (a) the interpreter's revealing of multiple perspectives and realities (Mishler, 1979; Noblit & Eaker, 1987; Spicer, 1976); (b) the mediator-broker's movement between social groups and structure (Greene, 1992; Noblit & Eaker, 1987; Torres, 1991); (c) the facilitator's involvement of participants and marginalized groups (Greene, 1992); (d) the consultant-educator's educative-reflective opportunities (Cousins & Earl, 1992; Torres, 1991); (e) the coach's drive to increase effectiveness/efficiency (Hendricks, 1993); (f) the investigator's search for hidden meanings.
(Smith, 1992); (g) the emancipator's discussion of ideological content (Greene, 1988; Noblit & Eaker, 1987); (h) the connoisseur's appraisal of the evaluative situation (Eisner, 1979); and (i) the social critic-critical ethnographer's challenge of ethical commitments (Greene, 1992; Thomas, 1993). Although the sentient is similar to these other roles, the synthesis results in a slightly different overall framework.

While challenging the social, political, personal structures and beliefs in the evaluation context, the evaluator, identified here as sentient, also simultaneously challenges her/his own structures and beliefs. S/he simultaneously reveals her/his own agendas, biases, beliefs, and understandings as s/he reveals those of the context. S/he detects, unmasks, and exposes "existing forms of beliefs that restrict or limit human freedom" (Neilsen, 1990). Closely aligned to cultural feminism, "which emphasizes a holistic and collective orientation to world and work experiences", the role of the sentient creates a kind of praxis, with continuous cycles of critique, knowledge construction, and social action (Hollingsworth, 1992, pp. 375, 398).

Cook (1991) poses the questions "What should the role of the evaluator be in an open, democratic society? Shouldn't cultural values and political action influence
decision as much or more than professional evaluative
feedback?" (p. 449). If, as House (in Alkin, 1990, pp.
31-32) suggests, the credibility of an evaluator is
essentially the fit of the "interest, ideology, and
background" of the evaluator with decision makers, and as
Law (in Alkin, 1990, p. 29) suggests one of the critical
factors of use is the credibility of the person doing the
program evaluation, then the role of the sentient may
have unique consideration for the future of the profes­sion of evaluation.

Ethical Considerations in Evaluation

Little has been written about the role of the sen­tient evaluator, and even less discussion has focused
specifically on the ethical situations encountered by
these evaluators in the evaluation process. This is
interesting if we agree with Smith, who states "evalua­tors probably encounter moral problems more frequently
than most people because they assume responsibility for
deciding what is good and of value in social programs and
policies" (1985, p. 5).

Most of the ethics discussion in the profession of
evaluation has focused on the moral issues of evaluation
(e.g., confidentiality of data, protection of human
subjects, etc.) and not on the moral issues of programs
and society (e.g., what are the values designed into the program? what do people operating the program value? what do people receiving the program value?) (Smith, 1985, p. 8). However, a salient theme emerging in the field of evaluation is that of ethical standards for the profession or at least a renewed discussion of the role of ethics in the field of evaluation. The role of the sentient may be reflected in some of the concerns behind creating "morally engaged" evaluation.

Schwandt (1992) believes that evaluation for too long has ignored the moral and ethical questions of engaging in evaluative inquiry, which has severely limited the ability of evaluators (and their clients) to reflect on "questions about the good life and the good society" (p. 136). Human inquiry cannot be divorced from "historical, moral-practical, and political locations" (p. 136), and politics and science are "ineluctably intertwined in evaluation practice" (p. 137).

We ignore the fact that what it means to describe, explain, or understand something about social and

---

6 An American Evaluation Association task force committee has produced a draft of guiding principles for evaluators (Newman, Scheirer, Shadish, & Wye, 1993). A tentative timeline for review and revision runs through January 1995. The Qualitative Research Special Interest Group of the American Education Research Association recently developed a casebook for teaching about ethical issue in qualitative research (Mathison, Ross, & Cornett, 1993).
political phenomena (which is what we are about when we do evaluation) issues from philosophical commit-
ments that control the way inquiry is conducted and
determine the significance that the resulting knowl-
dge will have (Shapiro, in Schwandt, 1989, p. 13).

"Those who conduct research within an evaluative context
are obligated to consider the moral implications that
their findings may carry for those who are affected by
the program in question" (Emihovich, 1992, p. 1). Evalua-
tion practice cannot ignore or escape from the necessity
of raising normative and ethical concerns embedded in
educational and social practice; evaluators must "ac-
knowledge to their audience their position on issues of
social justice" (Ericson, 1990, p. 19) by engaging in
self-conscious reflection of their values and articulat-
ing these values (Torres, 1991).

Schwandt suggests that "if there is truly a sense of
better living through evaluation, then we must have an
evaluation practice that is morally engaged, one that
addresses that core principles and values of individuals
and institutions" (1992, p. 139). In this setting, eval-
uation becomes a "public philosophy" that openly engages
clients, audiences, and stakeholders in dialogue about
"the moral meaning of practices", whether a program lives
up to the moral purposes claimed for it (pp. 139-140).
This type of "morally engaged" evaluation practice re-
quires (a) having a normative attention to individuals
and institutions, (b) approaching problems as concrete, grounded in a particular case or dilemma, and (c) engaging in social criticism (Schwandt, 1992).

Evaluators have to be self-critical, identify the moral issues and work to resolve moral problems in their own work, and behave in morally responsible ways (Smith, 1985). Being ethical in program evaluation is "a process of growth in understanding, perception, and creative problem-solving ability that respects the interests of individuals and society" (Sieber, 1980, p. 53).

Weiner, in her discussion of educational evaluation and social justice, echoes this notion of evaluation as deeply entrenched in its social, political, and ethical context; evaluation is shaped by the "subjects, events, and interpretations, that evaluators (and their sponsors) take to be important" (1990, p. 233). Personal and political values out of necessity enter into evaluators' work. In essence, evaluation should become a process of social inquiry where individuals understand their situation, where knowledge leads to emancipation, and where the focus is on program improvement, with the program grounded in the values of equality and liberation (Weiner, 1990).
For McGinty\(^7\), engagement in "morally engaged" evaluation and social enquiry as a feminist takes on particular ethical concerns:

1. Does the written report resonate with the actual lived experiences of those studied?
2. Is it useful to the subjects, enabling them to comprehend subordination?
3. Does the methodology lessen the divide between the [evaluator] and the [evaluated]?
4. Are the theoretical understandings and concepts accessible to those studied, or are they so abstract as to confuse those studied?
5. Does the written account demystify or clarify the underlying structural power relations that shape their everyday experiences?
6. Does the theory of the [evaluated] modify the [evaluator's] theory and interpretations?
7. Have I, as [evaluator], been an intellectual

\(^{7}\) McGinty in her discussion uses the terminology of research, researcher, researched, etc. The corresponding terms of evaluation, evaluator, evaluated, etc. have been substituted. Although the terms are not synonymous, McGinty's discussion fits within the framework of evaluation. (Research and evaluation are distinctly different endeavors - research focuses on investigation or experimentation aimed at discovery, interpretation, revision, and application of facts, theories or laws, while evaluation in some sense determines the worth or value of a project or program.)
tourist or voyeur, taking from the community without giving back to them? (1992, p. 6).

The researcher (evaluator) helps the respondents cope with their new knowledge, creates a situation without deception in which the researcher (evaluator) and researched (evaluated) are involved in the construction of reality, and looks to defining the "benefit" of the research (evaluation) to society not in terms of quoting other literature which says that the research (evaluation) needs to be done (typical of most Ph.D. dissertations!), but instead has "those to whom the benefits should flow, confirm the need or the conclusion" (McGinty, 1992, p. 16).

Similarly, Emihovich, in discussing feminist research and evaluation, quotes Seller (1988) "knowledge and politics are process rather than achievement, and that the commitment to engage in conversation to find out what the world is like is a moral and political commitment to a community" (pp. 180-181).

Communicating and Reporting Evaluation Findings

While factors about reporting/communicating evaluation findings have been addressed in the literature, there has been no systematic, comprehensive treatment of the topic. Most of the literature gives prescriptive
directions for organizing the format and content of evaluation reports and means of identifying audiences for communicating and reporting efforts; what has been addressed has not been in terms of the problems and issues currently faced by evaluators practicing in a variety of settings (e.g., internal/external, education, business, government, health care, and non-profit/foundations) and in a variety of roles (e.g., mediator, facilitator, educator, sentient, etc.).

Today, the concept of evaluation use is no longer just getting clients to "read" reports; it encompasses playing a joint role in constituting decisions, being cited in debates, being used in inservice training of professional, being used in educating future practitioners, and being used to reconceptualize social programs and problems. The agent of influence is no longer a single evaluation report presented to formal decision makers (Shadish, Cook, & Leviton, 1991. p. 450).

As Cronbach and colleagues (1980) argued decisions are rarely made, but are rather slipped into and are not based on some linear process of problem identification, information collection, information analysis, to information utilization.

Interestingly, many of the prescriptive directions given for increasing the use of evaluation reports/findings and increasing audience participation and ownership of results tie in very closely with the role of the sentient.
Involvement in Utilization of Findings

Brown and Braskamp (1980), in their summary of elements that enhance utilization of evaluation reports, remind us that "evaluators cannot wait until a report is written before they think about how evaluation information will be disseminated and utilized" (p. 92); evaluators must take an active role in utilization with continuous discussion, explanation, and collection of a wide range of information (Brown & Braskamp, 1980). Sentient evaluators are able to focus on change through open participation, negotiation, and discussion between groups; a sentient evaluator of a program is not passive and tries to collect a wide range of information to produce a holistic view.

Political Context

An evaluator must also be aware that politics affect the evaluation process, how information is communicated, and how that information is used. "An evaluator sensitized to the political context in which evaluation is taking place can facilitate maximum utilization of evaluation results" (Cahn, 1988, p. 1); this depends on the timing, audience, and format of the reporting efforts. Evaluation must be tailored to the political system, not applied as an abstract model (Cronbach, 1980). "The
evaluator has a political influence even when he does not aspire to it" (Cronbach, 1980, p. 67). A sentient evaluator has a clear picture of the political agenda of the program, the context, and her/his own beliefs. The sentient assumes a share of the responsibility for (a) development of a policy question; (b) translation of the policy question into an evaluation question that is researchable; (c) translation of the evaluation question into the evaluation design, performance, and reporting; (d) translation of the results into policy language; and (e) assisting in the use of the evaluation findings in the political process (Chelimsky, 1987).

Evaluator Credibility and Relationships

The sentient evaluator is sensitive to the audiences s/he serves, getting information to those with the greatest power to act on it, while also giving power to others and recognizing the social structure of power (Cronbach, 1980). The credibility of the evaluator greatly affects how information is received by the evaluation clients, audiences, and stakeholders; the evaluator must develop positive and productive working relationships and assist in reviewing evaluation findings and their implication for program improvement (Drezek, 1982). In order to produce useful evaluative recommendations, the evaluator
should (a) consider all aspects of the issue "fair game"; (b) draw possible recommendations from a wide variety of sources; (c) work closely with agency personnel throughout the process (collaboration keeps evaluation focused on relevant issues, involves personnel in developing recommendations, and reduces surprises, defensiveness, and disputes); (d) consider the larger context in which the recommendations must fit; (e) show the future implications of recommendations; and (f) stay involved after recommendations have been accepted (Hendricks & Handley, 1990).

The sentient evaluator tries to develop active relationships with clients, stakeholders, and audiences by assisting staff in developing findings and recommendations and recognizing their implications for program improvement. These groups become problem-solvers, framing evaluations findings "both positive and negative, as problem solving feedback" (Mokros, 1982, p.21).
CHAPTER III

METHODOLOGY AND DATA ANALYSIS

Participants

This study utilized indepth semi-structured and unstructured ethnographic interviewing techniques (Goetz & LeCompte, 1984; Spradley, 1979) on:

1. A purposive sample (n=12) of evaluators identified as sentient through review of: (a) their own statements in journal articles, professional papers, and/or published research in monographs/books related to viewpoints/behaviors exemplifying sentient evaluation; or (b) by their involvement in and presentations for the Qualitative Research, Women in Education, and Research on Evaluation Special Interest Groups of AERA on topics related to the sentient role; or (c) by their involvement in and presentations at American Evaluation Association Topical Interest Groups of Utilizing Evaluation and Qualitative Methods related to the sentient role; or (d) by recommendations by practicing evaluators of persons who exemplify the behaviors/perspective of a sentient.

2. A random sample of evaluators (n=7) who are members of the American Evaluation Association.

36
members of the American Evaluation Association. This random group of evaluators is used as a comparison to those evaluators specifically identified as "sentient" by the researcher. They serve as a means of contrasting "sentient" evaluators' practices and beliefs with the more currently common practice and belief in evaluation. The purpose of the comparison is to help generate hypotheses for future research, not to draw conclusions about predetermined hypotheses or generalize to other populations. The random group was chosen from the American Evaluation Association's membership directory.

Each potential participant was sent a personal letter outlining the study and asking for her/his participation (See Appendix B). Approximately ten days after sending the letters, the researcher contacted each person by phone to ask for their participation. Originally 20 sentient evaluators were identified as possible participants. Of these, 11 agreed to participate, four were

---

9 It should be noted that not all professional evaluators or persons who conduct evaluation activities in their occupations belong to the American Evaluation Association or the American Educational Research Association. Thus, the persons chosen for this study do not represent all members of the evaluation community or persons who conduct evaluation activities. Any conclusions drawn from this study must be framed within this limited context (members of AEA and/or AERA) of evaluators.
unable to participate due to time conflicts, and five were out of the country and/or could not be reached by phone or mail. One additional sentient evaluator was solicited for participation at the 1993 AEA conference in Dallas, TX following a paper presentation; this person agreed to be interviewed.

Twenty random participants were identified from the AEA membership directory. They were also sent a personal letter and contacted by phone. Only five of the 20 were able to participate, due to conflicts, time constraints, changes in job positions, and lack of interest. Another random sample of 10 names from the AEA directory was chosen and these individuals were contacted; two of the ten were able to participate.

Being concerned with only being able to arrange participation by seven random participants, as having unequal numbers might be considered a weakness in the study, the researcher looked at possible reasons why certain members of the random group chose not to participate, whether those that did participate are atypical of the random group, and if there are major differences between the random and target samples.

Looking at the educational level of the targets (the sentient evaluators) and comparing it to the 30 random names that were chosen from the AEA directory (thinking
that perhaps the targets might have more graduate degrees and having lived through the experience had some empathy/sympathy for the researcher and thus accepted out of kindness), proportionally the members of the target group with graduate degrees (master's, specialist's, or terminal degrees) is nearly the same as the members of the random group (target 18/20 = 90% and random 26/30 = 87% with graduate degrees). Of the seven persons in the random group who participated one has only a bachelor's degree, two are doctoral candidates, and the remaining four have terminal degrees (thus 86% have graduate degrees—typical of the entire contacted random group).

Then, the researcher considered the affiliation with higher educational institutions as a possible difference (thinking again that someone who actively participates in education research processes, i.e., master's and doctoral theses, would be more likely to agree to participate). Again, the difference between the two groups is minor (target 17/20 = 85% and random 23/30 = 76% affiliated). Of the seven persons in the random group who participated five (71%) are currently affiliated with a college or university (close in percentage to the entire contacted random group).

Gender was not a factor (target = 5 men out of nine contacted and 7 women out of 11 and random = 3 men out of
14 and 4 women out of 16).

Finally, the researcher considered that "interest" was probably the deciding factor; because members of the target group had written or presented on similar topics, they were probably more interested in conversing on the areas of focus of this study.

Inspite of the numerical difference in the random and target group, they are similar in a number of ways (see Demographics subsection in Findings), and thus this particular weakness is lessened.

A schedule of the interviews to take place at the annual AEA conference in Dallas, TX, November 1993 was developed in mid-October. One person was interviewed in October 1993 in Kalamazoo MI. Fourteen persons in total were interviewed over a four day period at AEA. Four persons did not attend AEA or had other scheduling conflicts. Two phone interviews were conducted with sentient evaluators and two with members of the random sample during December 1993 and January 1994.

Choice of Methodology

Ethnographic interview techniques and document analysis were appropriate methods for data collection in this study because its purpose was to explore the role of a sentient evaluator in order to expose the nuances,
complexities, idiosyncrasies, and context without manipulating or controlling the setting (Patton, 1987). These methods helped the research (a) get close enough to the interviewees to understand what they do and who they are; (b) capture what they say; (c) describe the interviewees, their activities, and their interactions; and (d) report direct quotations of what they say and write (Patton, 1982).

The interviews were a combination of three types of interview instrumentation: informal conversational, interview guided, and standardized open-ended (Patton, 1987). The standardized opened ended questions were used to gather basic background data on the evaluators. These questions were precisely worded, in a predetermined sequence so that the same type of information is gathered from each interviewee and at the same time in the interview process (See Appendix A for interview protocol).

The interview guided approach is more loosely structured, with topics and issues determined in advance, but with no particular order or wording of the questions. These questions were used to gather information about communicating and reporting, collaboration and audience/client/stakeholder involvement, and ethical dilemmas. These were more loosely structured to allow for a more natural flow to the interview process and so that
the interviewer can probe certain topic areas without interrupting or confusing the interviewee (See Appendix A).

Finally, the informal conversational style was used when questions and topics emerge from the immediate context of the interview process and need to be addressed or explored (See Appendix A).

**Development of Interview Protocol**

The development, piloting, and final revision of the interview protocols involved using the information gathered through literature review, past evaluation work and research, and feedback from practicing evaluators concerning the accuracy, relevancy, and appropriateness of the language of and issues in the questions.

Using (a) annotated bibliographies on communicating/reporting in evaluation, evaluation utilization, and internal/external evaluation; (b) papers, monographs, and texts on evaluation practice, evaluation theory, ethics in evaluation, evaluator roles, and feminist/marxist research; (c) previous personal evaluation reports, research, coursework, and papers on evaluation practice and theory; and (d) notes from discussions with practicing evaluators about communicating/reporting issues, evaluation utilization, and evaluator roles, the researcher
developed a preliminary interview protocol.

This protocol was distributed to four evaluators who have worked in a variety of areas, including education, economic development/public policy, business and industry, and non-profit organizations/foundations. Three have 15+ years of experience in evaluation, and one has 3+ years in evaluation. These four evaluators were solicited for feedback on (a) the relevance of the interview questions to the purpose of the study, (b) the clarity and focus of the questions, (c) the flow and logical order of the questions, and (d) the relevance and saliency of the issues raised in the questions to the field of evaluation. Their feedback was positive and encouraging concerning the clarity and relevancy of the protocol; none suggested major changes in the protocol's wording or order of the questions.

The following areas of inquiry were the major focus of the interview protocol: (a) model of evaluation used/method of inquiry or investigation; (b) role of stakeholders/clients/audiences in evaluation phases (level of involvement in development of evaluation plan, data collection and analysis, writing of "reports", and communicating and reporting findings); (c) communicating and reporting findings (perception of importance, individuals involved, methods used, and satisfaction with
methods); and (d) ethics and social responsibility (personal views/values, extent to which these enter into choice of evaluation projects, extent to which these enter into the evaluation process, and extent to which they are articulated to audiences/clients/stakeholders, etc).

Interview Process

The researcher used a portable tape-recorder with a microphone to audiotape the interviews within a closed setting (private room or office). Only the researcher and the interviewee were present during each interview session which lasted approximately 30-90 minutes in length. One interview was conducted in person during October, 1993 in Kalamazoo MI; 14 interviews were conducted in person at the annual meeting of the American Evaluation Association in November, 1993, Dallas TX; and four persons were interviewed by phone. All but one interview was audiotape recorded and transcribed. One interview could not be taped (due to technical difficulties with a speaker phone) and was recreated through interview notes.

Although the interviews actually took on two forms, in person and over the telephone, the resulting data is the same. It was expected that the interviews conducted in person might be lengthier, have more detail and depth,
and have a crisper flow of ideas due to the face to face contact and the ability of the interviewer and the interviewee to play off of each others gestures. This, however, is not the case; the interviews that were conducted over the phone when reviewed through interview notes (see "site analysis" form in Appendix E) and the transcripts themselves bear no striking difference with those interviews conducted in person. This is probably due to the fact that each person interviewed received an overview of the study which listed the main areas of focus for the interview; so in a sense, each person, regardless of how they were interviewed, had the same information on which to prepare for the interview.

The only transcript which may be slightly different is the one that had to be prepared from intensive note-taking (due to the malfunction of a speaker phone prior to beginning the interview) versus being transcribed from an audiotape. The main difference between this transcript and the others is that it lacks natural speech patterns, with all of the gaps, fragments, repetitions, and awkward sentence structure used in conversations. The researcher tried very deliberately to capture this person's vocabulary/language, her/his emphasis/focus on certain topics and ideas, and of course had her/him review the transcript, as did all the other participants, but it still
is slightly different from the other transcripts. However, the difference did not seem to make any difference in the coding or analysis of its data. When reading the coded outputs, one cannot distinguish this person's data from the other 18 participants.

Each person received a transcript of her/his interview and was asked to make any changes or additions that s/he felt were important for clarity. Thirteen participants returned their transcripts with alterations, ranging from incorrect names/dates to additional comments and reflections concerning the interview questions. These transcripts were corrected and used for data analysis.

All tapes produced during the interviews are the property of the researcher and have been only in her possession before/during/after the interview sessions and research process. The tapes are identified by the date of the interview, the setting, and the interview identification number (ID#).

The identification number is known only to the researcher and the interviewee so that transcripts and interpretive descriptions can be given to the individual interviewees for feedback and revision. The code is a three-digit number (e.g.; 003) assigned to the interviewee by the researcher and secured in the researcher's personal files so that only she can match the ID# to the
interviewee's name.

Although all efforts have been made to assure the confidentiality of the information, it should be noted that it is not anticipated that the information collected during this research project will be "controversial or sensitive" or in any way will jeopardize the well-being, livelihood, or professional standing of the participants. Each participant was given an informed consent form which was signed and submitted to the researcher prior to the interview process (See Appendix C).

Data Analysis Overview

The interview transcripts were analyzed using an emergent design, in which a series of reflective and evolving data analysis phases took place in order to capture the varying themes, categories, and/or "stories" that emerged from the data (Lincoln & Guba, 1985; Tesch, 1990; Wolcott, 1990).

The analysis of the interview transcripts followed the basic design presented below in Figure 1.

Transcription

The researcher completed all transcription of the interview tapes within a closed setting (her personal home office) so that no other individuals had access to
the "raw" information (the names of the participants, the tapes, or the transcribed data). Any names or identifying information discussed on the tapes are substituted for using the interviewee ID# or noted using dash marks or pseudonyms (e.g.; Western Michigan University = ---- University).

Interview --> Transcription --> Preliminary reading and identification of themes --> Development of preliminary coding scheme --> Use of coding scheme on transcripts or documents --> Analysis of coded data to identify themes or categories --> Review and revision of coding scheme --> Use of coding scheme on transcripts or documents --> Analysis of coded data --> Development of categories and themes --> Development or writing of "stories" or descriptions --> Distribution of "stories"/descriptions to interviewees for 'member checks' -->

Final "stories" on communicating/reporting, collaboration/relationships/involvement, ethical considerations, and other categories that emerge.

Figure 1. Data Analysis Process.

A printed copy of her/his transcript was sent to the individual interviewee. The purpose of providing the interviewees with copies of their transcripts and drafts
of analysis description or stories was to have them review, revise, confirm, or reinterpret the information.

If necessary, the researcher and the interviewees would have discussed any major discrepancies or disagreements in the analysis and interpretation of the information or in the transcripts themselves, with the interviewee having the final authority. The interviewee was able to withdraw his/her participation in the project no later than 2 weeks after the final "member checking process" by submitting a written request to the researcher and her doctoral committee (See Appendix C for copy of consent form/research and data agreement). If necessary, after that time all efforts would have been made to negotiate any final discrepancies or disagreements between the researchers and the participants. These precautions did not come into play in this study.

The transcriptions are housed on computer discs accessible only to the researcher and identified only by the interviewee identification number. The transcripts are stored in this format to facilitate the data analysis using ETHNOGRAPH (Qualis Research Associates, 1988), a qualitative text-analysis software package.

Coding Scheme

The first level of coding was developed from the
major topical areas covered in the interview protocol and any addition topics discussed during the individual interviews. After reading the transcripts twice, a list was made which covered the topical areas in the interview questions and any additional topics mentioned in the transcripts. Twenty-one codes were developed in this first level as presented below in Figure 2. The codes are five and six letter abbreviations for the topic areas (the abbreviations are needed for inputting this information into the computer text-analysis program ETHNOGRAPH).

Using these codes, each transcript was read through and coded twice. The first trial of coding was a trial run to get the researcher used to using the codes and refamiliarized with the transcripts. The first coding was done by writing in pencil the code for the designated line next to the line, using brackets to designate where a segment of lines had the same code. The second trial of coding took place two days later so that the researcher could approach the second trial without remembering by rote how she had coded the transcripts in the first trial. The transcripts were recoded using the same transcript that was marked during the first trial. The codes were positioned to the right of the transcript data, so the researcher was able to re-read the transcript narrative without seeing the first-trial codes. Second-trial
codes were recorded in blue ink to distinguish these from the first-trial codes.

SETTING = Settings/areas conducted/conducting evaluation
PRFROL = Professional role(s)
YEARS = Years experience conducting evaluations
PROCES = Typical evaluation process/inquiry method
CLCOM = Client communication
PURPOS = Purpose of evaluation
BCKVAL = Background/values
SOCRES = Social Responsibility/Role in Society
COLLEA = Colleagues
CLIENT = Client/stakeholders/audiences
CLTINV = Client/s/a involvement
COMREP = Communicating/Reporting general
CRIMPT = Com/Rep importance
CRSATF = Com/Rep satisfaction
CRACTV = Com/Rep activities
CRFDBK = Com/Rep Feedback
CHOICE = Choice of evaluation clients/contracts/experiences
ETHDIL = Ethical dilemmas/conflicts
ETHRES = Ethical dilemmas resolutions
FTISSU = Future issues in evaluation
FTACTV = Future activities to address issues

Figure 2. Coding Scheme - First Round.

After these two trials of coding, the transcripts were read a third time to see where any codes were redundant and to find if particular codes had not been used at all. From this coding/review cycle two codes were eliminated because they were redundant with other, more specific codes (CLIENT and COMREP). No narrative segments lacked codes and no topical areas were found that necessitated the creation of any new codes.

During this second coding/review cycle the remaining 19 codes were "grouped" together in terms of similar
topical area or overlap of influence as presented in Figure 3. This grouping of codes facilitated the use of the ETHNOGRAPH software. (Appendix H contains a "code book" which outlines a clear definition or descriptor for each code and where individual codes relate to particular interview questions.)

SETTING = Settings/areas conducted/conducting evaluation
PRFROL = Professional role(s)
YEARS = Years experience conducting evaluations

-------------- Group I

PROCES = Typical evaluation process/inquiry method
CLCOM = Client communication
CLTINV = Client/s/a involvement

---------- Group II

PURPOS = Purpose of evaluation
BCKVAL = Background/values

------------ Group III

SOCRES = Social Responsibility/Role in Society
COLLEA = Colleagues

---------------- Group IV

CRIMPT = Com/Rep importance
CRSATF = Com/Rep satisfaction
CRACTV = Com/Rep activities
CRFDBK = Com/Rep Feedback

--------------- Group V

CHOICE = Choice of evaluation clients/contracts/experiences
ETHDIL = Ethical dilemmas/conflicts
ETHRES = Ethical dilemmas resolutions

--------------- — Group VI

FTISSU = Future issues in evaluation
FTACTV = Future activities to address issues

--------------- Group VII

Figure 3. Coding Scheme - Second Round.
The ETHNOGRAPH software allows the researcher to search via the computer for up to 99 codes simultaneously, either in the order designated by the researcher or by alphabetical order. Searching the codes in their predetermined "groupings" allows the researcher to more easily organize the computer printouts and facilitates the reading of the data and the development of the findings, since similar topics/issues are near each other on the computer printouts, eliminating extra paper shuffling/sorting or the need to physically cut and paste the printouts.

Once all of the transcripts were recoded using the remaining 19 codes, these codes were entered into ETHNOGRAPH. This computer program enables the researcher to code individual lines or a series of lines in a document and then search each document (interview transcripts) using these codes.

Seven searches were run on the transcripts, utilizing ETHNOGRAPH's ability to search codes in the order determined by the researcher, based on the seven "grouping" of the codes.

Following the searches, within each grouping those responses/data that belonged to the target ("sentient") evaluators and those that belonged to the random group were separated into two subsets.
Analysis of Coded Data

Following the seven searches run on the transcripts, each "grouping" was put in a separate file folder. Each folder was read for consistency of the narrative information with the designated code (i.e., Were any lines or series of lines miscoded? Was any information confusing because it was taken out of context? or because pieces of the line was lost?, etc.). No problems were found within any of the coded outputs.

One feature of ETHNOGRAPH that is particularly helpful during analysis is that for each search the computer software produces a list of those documents (interview transcripts) in which each code was not found. This facilitates going back to individual transcripts to see if there is indeed no data that matches that code or if such information was just overlooked by the researcher. On two codes the researcher found areas where information had been overlooked or listed under another code (PURPOS and ETHDIL). These portions of the transcripts were recoded and added to the respective file folders.

Once all of the coded outputs were read for consistency and accuracy, for each "grouping" folder, a list of basic issues/statements from each respondent was developed on separate sheets of paper, one sheet for the target group, one for the random group per folder. From
these basic lists, the researcher looked at the similarities/differences between individual evaluators and the two groups (target and random) and began to brainstorm ways to present this information: Issue-lists/Areas of Focus Descriptions, on Continuums, as Contrasts-Counterpoints, as Composite Stories or overviews. [Although it may seem premature to think about how the data should be presented when beginning the analysis of data, it is vital to simultaneously think about the meaning of the data and how that "meaning" can best be conveyed].

Each "grouping" folder was read through a minimum of three times, in varying order, to try to capture the many nuances of the topics/issues, similarities/differences, relationships, consistencies/inconsistencies in the data. For each reading of each folder, detailed notes were taken on what the data seemed to suggest and ways of presenting the "meaning" of the data.

One phase of the data analysis that was peripheral to the research process, yet vital, was the researcher's preparation of a paper presentation for the American Educational Research Association's annual conference April 1994 concerning one subpart of this overall research process. In this paper, the major topics discussed by the target group are presented in terms of their impact on the future of the profession of evalua-
tion. Preparing this paper was extremely helpful in developing additional ways of presenting the findings and clarifying the researcher's focus on the topical areas of (a) the model/method of inquiry used, (b) the role of stakeholders in the evaluation process, (c) the purpose of evaluation, and (d) communicating and reporting evaluation findings (Piontek, 1994).

The findings are presented in a variety of ways to best capture the meaning of the data, the relationships between the target and random groups, the similarities/differences between individual participants, and the overall composition of the sentient evaluator -- the inconsistencies, variances, and nuances.

Credibility, Transferability, Dependability, and Confirmability of the Study

The areas of reliability, validity (internal and external), and objectivity are addressed in naturalistic inquiry in terms of credibility, transferability, dependability, and confirmability (Guba, 1981).

Credibility is when multiple constructions of reality are adequately represented and the reconstructions of these realities are deemed credible to the interviewees/informants (Lincoln & Guba, 1985). The study attempts to satisfy the criteria of credibility (a) by conducting 'member checks' with the interviewees in which they react

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
to, revise, and perhaps reinterpret "stories" created from the data; and (b) by debriefing (discussing) with practicing evaluators outside of the study to test emerging themes and categories (No identifying information or "raw" data-transcripts was accessible to these persons. Only interpretative ideas or emerging themes or codes were discussed for reactions to their plausibility and credibility). (See Appendix D for notes on research process and contacts with practicing evaluators).

A "member checking" format was used to help obtain feedback from the interviewees on the credibility and accuracy of the "stories" or descriptions produced by the data analysis and writing sessions (See Appendix F).

Draft findings were distributed to each participant so that feedback concerning the accuracy, consistency, and value of the findings could be solicited. Each of the six subsections that make up the findings portion of this study were checked by five participants, made up of both the target and random group. Each participant was given one or two subsections to check (i.e., one person was given the The Purpose of Evaluation and Role of Stakeholders/Clients/Audiences in Evaluation Process sections, while another received The Model/Method of Inquiry Used and Sentient Evaluators, in all there were 10 combinations).
Since the numerical codes given to each participant were altered in the presentation of the findings, the researcher did not deliberately try to send an individual participant his/her direct quotes that were presented in the findings\(^9\).

No personal or controversial information was presented in the direct quotes so it is impossible for the participants to identify one another -- in fact since the code given to each participant in the presentation of the findings are re-numbered an individual may or may not recognize his/her own quotes. The emphasis here is to get feedback concerning the accuracy, consistency, and value of the findings as presented and the meaning(s) ascribed to the data, not on whether an individual can recognize her/his own words.

This feedback influenced the findings by (a) assisting the researcher in recognizing some of the nuances in the meaning of certain quotes (especially the sections on the Purpose of Evaluation, Ethics and Social Responsibility, and Implications); (b) offering suggestions on how

---

\(^9\)The numerical codes given to each participant in the interview/transcription process were altered in the presentation of the findings in order to make it clear which participants are from the target (sentient) group and which are from the random group. Those in the random group are designated as R01 through R07, the target group T01 through T12.
to preface the use of continuums and what they represent; and (c) highlighting language/vocabulary used by the researcher that may be misleading in terms of suggesting generalizability (or similar concepts) where this was not the intent.

Transferability is the process of "thickly" describing the time and context of the inquiry (Lincoln & Guba, 1985). The study attempts to describe the role of the sentient evaluator in depth and in detail by using authentic data from the interview transcripts.

Dependability is the process of demonstrating credibility of the process of inquiry and of the findings. Detailed notes and descriptions of the development of the interview protocols, the interviews (context, format, content), the data analysis procedures, the writing process and members checks, and any conversations or debriefing sessions with practicing evaluators, colleagues or fellow students, and committee members have been developed (See Appendix D).

As a means of recording main themes, impressions, speculations, and implications evolving from the interviews, a modified "site analysis" form was attached to each interview transcript. This was useful in tracking the credibility, dependability, and confirmability of the study (See Appendix E).
Confirmability is the degree to which the findings of the inquiry grow out of the subjects and the conditions and not by the bias of the researcher. The study provides a detailed discussion of the researcher's personal and professional background and develops detailed notes and descriptions of all phases of the research (See Appendix D).
CHAPTER IV

PRESENTATION OF DATA AND FINDINGS

Overview

The final presentations of data and findings generated in this study take several forms\textsuperscript{10}. The findings are presented in a variety of ways to best capture the meaning of the data, the relationships between the target and random groups, and the similarities/differences between individual participants.

Continuums for contrasting the evaluators with each other emphasize the major areas that grew out of the data collection and analysis: (a) the purpose of evaluation, (b) the involvement of various groups in the evaluation process, (c) ethical dilemmas faced by evaluators, and (d) the extent to which evaluation has social responsibilities.

The role of the sentient is also contrasted with the random interviewees in terms of their communicating and

\textsuperscript{10}The numerical codes given to each participant in the interview/transcription process were altered in the presentation of the findings in order to make it clear which participants are from the target (sentient) group and which are from the random group. Those in the random group are designated as R01 through R07, the target group T01 through T12.
reporting methods and the most salient issues for the evaluation profession. These findings highlight the main similarities and differences in these two areas by presenting direct quotations from the participants.

Demographic Information

Of the twelve targeted persons ("sentient" evaluators) who participated, four have over 20 years of experience in the field of evaluation (the highest being 27 years), five have 15-19 years of experience, and three have 10-14 years. Educational evaluation is the main area of focus for this group (ranging from local district programs to nation-wide initiatives, from traditional K-12 settings to non-traditional settings of environmental education policy, continuing education, and museum education). Although educational evaluation is the main area of focus for this group, some have evaluation experience in health care (2), human-social services (3), or business-industry settings (3). Five are men; seven are women. One person has only a master's degree (8%), one is a doctoral candidate (8%), and the remaining 10 have terminal degrees (84%). Thus, 100% of the target group have a graduate level degree.

Of the seven random persons who participated three have over 20 years of experience in the field of evalua-
tion, one has 15-19 years of experience, and three have less than 5 years. Two practice primarily in non-profit health and human service organizations, one in non-traditional education settings, three in K-12/higher education, and one in criminal justice. Three are men; four are women. One person has a bachelor's degree (14%), two are doctoral candidates (28%), and the remaining 4 have terminal degrees (58%). Thus, 86% of the random group have a graduate degree. The one member of the random group who has a bachelor's degree works as a team with an evaluator with a Ph.D., and thus was very familiar with many of the theoretical underpinnings of evaluation.

The members of the target and random groups live in all parts of the U.S.: target = California, Colorado, Florida, Illinois, Indiana, Kansas, Maryland, Minnesota, Ohio, and New York (2); random = Arizona, Connecticut, New York, Texas, Virginia (2), and Wisconsin. One member of the target group is from Israel, received her/his Ph.D. in the U.S., and has practiced evaluation and conducted research in the U.S..

Although the researcher designated these persons as "evaluators", they were specifically asked what they saw as their main professional role. Six of the targets identify themselves as evaluators/evaluation specialists, and three identify themselves as university faculty.
members/educators. One person uses the title "innovations facilitator", one person refers to her/his role as generalist/action researcher, and one person uses the term "field person".

In the random group, three identify themselves as evaluators/evaluation specialists, and one identifies herself/himself as a university faculty member/educator. One person refers to her/his role as a researcher-technologist, one person is a graduate student, and one person uses the term "field person".

The Purpose of Evaluation

The participants were asked what they saw as the main purpose of evaluation. Each person had a unique answer concerning this question, but the responses can be loosely sorted into six categories:

1. Meeting consumer/stakeholder needs (R07, T06, T07, and T11):

And so I try to get them off center about validating or not validating something that they do, but trying to decide who they are trying to serve and how they can do that better.R07

---

Note that each participant may actually fall into many of these categories; it is difficult to try to improve a program without also increasing understanding, and/or meeting decision-maker and stakeholder needs. In sorting the participants into these categories the researcher focused on the main purpose discussed by each participant.
I think that we ought to be trying hard to understand the values and situations of our clients and take on their concerns at least for a while.T06

The purpose of evaluation, as the purpose of research too, is unabashedly reformist. I work in schools and I am concerned with programming for children, conditions in which children live and work in schools, so I guess that underlying all evaluations is the notion that we ought to provide better, just, appropriate programs for kids...ultimately if it doesn't make life for kids better, then there isn't any reason to do it.T07

Is it meeting the needs of who it is intended for. That is the main purpose. A program may be well run, but not doing anything for the people.T11

2. Meeting decision-maker needs (R01):

Making judgements about programs, but more precisely or more practically, bringing information to bear on decisions about programs. I mean it is a real decision-oriented, management-oriented approach to evaluation.R01

3. Program improvement (R02, R03, R05, and R06):

It is to try to find out how to develop something to a higher plateau.R02

Contrary to what some people think it is not to scare people to death!...I think that evaluation can be a very good tracking system to alert people in the program to problem areas they have, and it is also--there are strengths within the program not even recognized by the staff, and being able to point out some of these strengths in the program helps build a stronger foundation, helps give the staff pride in what they are doing, and really helps to build an alliance between evaluators and staff.R03

To improve programs or products.R05

The main purposes of evaluation if we speak generically are to provide answers to questions for people who are involved, if it is a program evaluation, the questions people in the program have and how the work that they are doing...how they can do it more
effectively.

4. Equity/justice (T03, T04, T05, T08, and T10):

In empowerment it is self-determination.

Evaluation must promote the democratic aims and claims of equity and justice.

I like the goal of liberation...being a vehicle to agency.

Stop oppression! I think we can be very helpful in improving the quality of life.

I see it as a matter of social critique...critical in a critical theory kind of way...to say -- does this make sense? is it fair? is it just? whose interests are being served? is this suppression necessary? does it matter?

5. Determining merit/worth (R04 and T02):

Accountability...is this program worth the money.

I see it as a matter of determining the merit or worth of something...determination of merit or worth or value of something.

6. Increasing knowledge/understanding (T01, T09, and T12):

To get a better understanding of the certain situations that you want to do something with or think about. I would think that the key phrases would be 'understanding' and 'practical situation'...it is understanding of issues and problems that you want to do something about.

For myself, it is to make sense of programs, it's to help people understand what's happening in order to improve, what is going on...it is empowering people with knowledge.

I think that it is to facilitate the client's, for lack of a better work, the client's self-understanding and their own ability to make sense of and evaluate what they are doing. I think that our job as evaluators is to help them engage in some kind of
critical self-reflection...I really think evaluators should be more like teachers.\textsuperscript{12}

Again, emphasizing that each participant may actually fall into many of these categories, another way of depicting the various purposes to evaluation discussed by the participants is to think in terms of a continuum\textsuperscript{12}. A continuum whose poles represent the level of detail of the purpose, ranging from a macro-global-societal focus to micro-program-immediate stakeholders focus, as presented in Figure 4.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4}
\caption{Purpose of Evaluation.}
\end{figure}

Notice that in their discussion of the purpose of evaluation, there is some distinction between the target group and the random group. The random group clusters around the micro-focus, while only members of the target fall into the macro-focus. Certainly, this is not a

\textsuperscript{12}These continuums are artificial in the sense that the two poles that are presented may not really be diametric opposites in terms of the evaluation philosophy/viewpoint they fall under or their appropriateness/utility in the evaluation process. The purpose here is not to represent one end of the continuum as "good" and the other end as "bad." The continuums simply provide a means to present similarities/differences in a graphic form.
"black-and-white" dichotomy. Other members of the target group fall all along the continuum, but only members of the target group clearly focused on a more global, societal-level equity and justice purpose to evaluation.

The Model/Method of Inquiry Used

The participants were asked to describe a typical evaluation process or method of inquiry that they frequently use in their evaluation practice. Most responded that to some extent the method/model is dependent upon the context and the program under investigation, but there are some "general tendencies" in their evaluation practice.

For the random group the emphasis is on

1. Being involved from the beginning of the evaluation process and in contact with key decision makers.

The first things I always suggest to people is that I should be involved as an evaluator as the project unfolds, I'd like to tell people to involve me from the beginning in terms of program design. And when I am involved from that perspective, I think that it is much more meaningful to the people I'm working with...I always try to do a sort of a typical approach, decision-maker. R01

We go to management to make formal reports, in writing, orally, and get feedback.R02

Slowly they have taken the advice to have me involved when they are beginning a program.R05

2. Using a mix of qualitative and quantitative meth-
ods, borrowing from different models.

We use a basic mix of qualitative and quantitative.

R03

Constructivist-hermeneutic types of things...either quasi-experiment or experimental design to follow the impact, quantitative type stuff. R04

I tend to use a model that is more, I guess you would call it eclectic in that I tend to borrow from several different models in doing an evaluation.R07

3. Eliciting participation from program staff and key stakeholders, especially the "front-liners" who have direct contact with clients.

What I try to do is identify stakeholders, other people who have some other influence or interest in the program and I try to sit down with them and talk about what are the questions.R01

The approach that we use is to identify these different groups and get their input.R02

We've elicited a lot of participation from the program staff because there are a lot of times when they have been able to tell us where to look for data in places that aren't necessarily obvious...it is important to have the buy-in of everybody...because you never know where you are going to find the most amazing things.R03

We have to be responsive to parents because the parents are very involved. We have to be involved with the teachers, we have to be involved with the faculty members, the university administration...so everyone is very involved in it.R05

It always involves a lot of involvement with the client and at least to some degree with the service providers, the 'front-line folks'...it varies a great deal but I always strive to involve up front the people who will be, the stakeholders, the people who are most affected.R06

4. Using a variety of data collection and communica-
tion methods, including observations, surveys, discussions, and workshops.

A senior faculty member goes out and observes the classes, talks to teachers, our administrators go out and do the same thing. We conduct workshops for the teachers...collecting information from the students is more of a survey base, it is a mixed-bag to get the type of information that we need.R05

I use as many different strategies as I can to get at what is going on in a program.R07

5. Looking at the context and process, as well as the outcomes.

It depends on whether we want to look at a process-implementation or an impact. Sometimes we do both, sometimes we do one or the other.R04

I try not to limit myself to only those things that might be the stated goals and objectives of a given program...I've always been concerned with not looking just at product outcomes, but also context and process.R07

For the target group the emphasis is on

1. Studying the field in which the evaluation is to be conducted (through people and literature).

Studying the field in which the evaluation is to be conducted and studying through literature or by meeting other people.T01

2. Involving varied, diverse clients/stakeholders to develop questions/areas of focus for the evaluation (meetings, surveys, focus groups).

I am committed to as participatory a process as I can do...however, it is very difficult and evaluation is not a context that is very well matched to a participatory value framework...I spend a lot of time to get very diverse views on the agenda, about what questions should be asked. I can do that.
through meetings, I have interviewed as many as 50 different stakeholders at the beginning of a process to see what their views are...it is agenda setting and I do think as much as possible having a lot of people contribute to that.T04

I would characterize my practice now as being far more focused on collaboration...notions of genuine participation. T07

My approach is to be as open as possible and to bring in people and sit down and discuss and keep the dialogue going...brainstorming on what evaluation means to you...we talked about what kind of questions made sense based on the program. We would meet before we went off to do our respective things each day...we would go over the evaluation results.T08

It is collaborative, the type that I am interested in is collaborative action research...many times evaluations are not necessarily prescriptively collaborative, where as I am saying that's how you should do it.T09

I meet with the people who are going to pay me first, and secondly right off of the bat I conduct a series of focus groups with beneficiaries, and other stakeholders who are pertinent, and really try to flesh out from them...what I do say is what I need to hear is what are the issues that you feel are relevant for somebody coming in and looking at the program, and would you want to know more about and would like to speak to yourself.T10

3. Using ethnographic techniques to provide context and meaning for the evaluation, using an anthropological viewpoint and case study reporting methods.

One would be an ethnographic focus where you are using ethnographic concepts and techniques, and applying them to evaluation, which ranges from participant evaluation where you are staying with folks for a long time, listening for the emic or insider's perspective of reality and to get their view of whatever is the context of your work. Ethnographic techniques provide some context to evaluation so you can understand it better.T03
I guess that I'm inclined to do something that you would call case study evaluation, that takes some kind of, Stake calls its an 'anthropological perspective' and my sensitivities lie in that direction.

I've been influenced in particular by Bob Stake and his responsive evaluation and Mike Patton and his work on qualitative research, also feminist evaluation and evaluation as social justice.

4. Maintaining flexible, responsive evaluation activities (using multiple methods) and feedback.

I really feel evaluators should be able to employ several different methods of data collection methodologies that would be suited to the particular program or policy or whatever is being reviewed...we should be able to adjust ourselves methodologically, using at least a few different kinds of methods.

Responsive evaluation...I provide them with a considerable portfolio of previous work done and urge them to see if that is the kind of work they want done.

We kept the evaluation plan open, but with enough parameters that we could move forward and start data collection...it evolved that the instructors asked for daily feedback...we made up a questionnaire to use everyday.

5. Assisting clients/stakeholders in internalizing/institutionalizing evaluation as part of their planning and management.

Empowerment evaluation is much more focused on using evaluation so that it is internalized in a program and becomes part of their planning and management rather than an add-on parasite kind of thing. It is aimed toward helping people become more self-determined and in control of their own lives, by establishing goals and strategies.

The fact is that they have it and I never had the power in the first place...you are letting people see that they have the authority to take some action.
even if it is to disagree with me as the evaluator. You are just letting them see that they can do this, make a change, be involved, take action!T05

What I have tried to think about is what does it mean for all of us to participate in some common endeavor and to think about defining that and that is really recast in some significant ways what we have done, so that those folks are in fact engaged in thinking about what is important and to know what kinds of questions to ask, how we should collect data, what we should do in terms of writing a report.T07

The process that I am very excited about is one whereby I teach teams of people, typically from a school building, the action research process. Which is to say that they frame a problem, something in their own work that is bothering them, and then they design a plan that they believe will fix that problem, they take the action, they actually act, collect data themselves and reflect on what happened by looking at the plan, looking at the data, and then that reflection leads cyclically back to another cycle.T09

6. Being "planted in the community", continuous long-term relationships with program participants, being involved in the program so as to experience the same issues facing the program deliverers and recipients.

I really think that you have to be planted in the community.T05

We typically understand participation-evaluation to mean—you know you folks who are running the program or the recipient of the program, you are welcome to come over and pretend that you are the evaluator for a while—what I have tried to think about is what does it mean for all of us to participate in some common endeavor and to think about defining that and that is really recast in some significant ways what we have done, so that those folks are in fact engaged in thinking about what is important and to know what kinds of questions to ask, how we should collect data, what we should do in terms of writing a report. But also having some reciprocity in that
having the evaluators as participants in the program itself...it is this notion of equalizing the relationships in the context of the evaluation...and also adds a dimension to the evaluators' understanding of the program. If you get in the position of having to deal with the issues that program deliverers or recipients get put in, then you have different sources of data about the program and whether or not it is good or bad.T07

I met with them regularly...and they said, it got to be funny, they called me a 'friendly nag'...and so I tried to go out regularly enough so that kept the process going...the second year, they saw that they didn't have to do as much, they just had to do it regularly, and so they really learned that...they called me the 'cheerleader'...it wasn't really my technical skills, well sometimes it was...but they really oftentimes needed somebody to say you are doing fine...now this year they have integrated it.T09

7. Not just setting up a context of rich description, but also making evaluative statements -- these statements are developed jointly by the evaluator and the clients/stakeholders.

I will say here are the key findings and then we will have the discussion, for two purposes. One to find out if it makes sense and the 'so what?' questions. I will try to do that including as many persons as possible.T04

Ultimately you have to make an evaluative statement. You have to draw an evaluative conclusion...one of the things that I have been trying to do is to still take that responsibility to heart, but not have the evaluator do it alone. I think that you can still do that in a very serious, reflective way and be honest about it and that a far larger group than simply the evaluator can be involved in rendering those conclusions.T07

Well you can imagine an evaluator saying 'well we need this'...we said instead 'this is a comment that was given to use, can you tell us what the situation is?' Now that is a completely different understand-

I can make a real contribution there and see real changes and do some good about it, but also on a larger scale watching my life go by, I want to see that I have done something that matters to me, that matters in my value system.

Advocacy is what happens before illumination...what really happens is that you are listening to people talk...you are having some kind of conversation, a text, with somebody that you don't really know and they are going to say things that they had not thought they thought...and things that they had certainly not heard themselves say...so it is almost like an analysis process...to let them get behind what they are thinking...you end up communicating your biases as a person on the planet.

You know you honestly look at yourself and what it means to be a human being and what gains you entree and operating in another culture...it was one of those things of a feeling of real human connection and I guess that I feel like it is essential in evaluation that you bring that out.

Not surprisingly, many of these practices discussed by both groups coincide with basic evaluation tenets: involving clients/stakeholders, establishing communication/feedback mechanisms, being flexible and responsive to the unique nature of the program and its context, and making an "evaluative statement" about the merit or worth of a program. These ideas are not unique; they have become standard practice in most evaluation settings. For instance, as stated in The Program Evaluation Standards (Joint Committee, 1994), Stakeholder Identification (Utility Standard 1) suggests that the evaluator contact...
representatives of stakeholder groups to get their perspectives and involve clients/stakeholders in designing and conducting the evaluation. Program Documentation (Accuracy Standard 1), Context Analysis (Accuracy Standard 2), and Described Purposes and Procedures (Accuracy Standard 3) promote having the client/stakeholders describe the program being evaluated, discussing their ideas of the purposes and uses of the evaluation, and thoroughly describing the context of the program/evaluation. Similarly, Justified Conclusions (Accuracy Standard 10), Values Identification (Utility Standard 4), Evaluation Impact (Utility Standard 7), and Political Viability (Feasibility Standard 2) focus on communication evaluation information, interpreting evaluation data/findings, and soliciting feedback on data/findings from clients/stakeholders (Joint Committee, 1994).

What differs in the target group from standard practice and the random group is the extent and focus of the evaluator's participation in a program setting, the role s/he plays in developing an institutionalized evaluation "ethic," and the promotion of human advocacy or empowerment or liberation of individuals (Fisher, 1993; Gray, 1993; Moe, 1993). As Fisher discusses in Independent Sector's report on evaluation, the sentients would
agree that the evaluator is "responsible for conducting an evaluation process that makes every effort to enhance the effectiveness of the organization by creating an environment conducive to learning and change" (1993, p. 33).

Many of these "sentient" evaluators see the purpose of evaluation as being somehow connected not only to developing a clearer understanding of a program and establishing its merit/worth and how to improve it while assisting decision-makers, but also as being connected to social critique. Evaluation may be a means of creating a social critique of what is just and fair in society, looking at whose interests are being suppressed, and facilitating clients' self-understanding and reflection on what they are doing. This critique may lead to reform, empowerment, or liberation.

The Role of Stakeholders/Clients/Audiences in Evaluation Process

After describing their method(s) of inquiry, the participants were asked to further clarify how they specifically involve clients/stakeholders in their evaluation process, from developing an evaluation plan, to data collection and analysis, from writing of "reports" to communicating and reporting their findings.

Again, the extent of involvement by cli-
ents/stakeholders is dependent upon the context and the program under investigation, as well as the length of the evaluation process. The participants from both groups stress the importance of client/stakeholder involvement in the development of the focus of the evaluation and the evaluation questions to be explored. Often, clients/stakeholders are involved in critiquing and interpreting draft reports or findings.

Although client/stakeholder participation is valued, especially at the "front and back ends of the evaluation", some of the target evaluators mentioned that it is often necessarily limited participation. They mentioned failed attempts to involve clients/stakeholders in the development and selection of methodology, the logistical difficulties in getting feedback from diverse groups, overwhelming clients/stakeholders with "unfiltered" data which they could not interpret, and burdening clients/stakeholders with writing tasks (case study drafts) with which they have no comfort or expertise.

Sometimes it does not seems possible to me to design an evaluation and then go show it to stakeholders groups, because either the situation is not right for doing that--it is too difficult to do logistically, and/or the people don't really care or are not able to respond to the design because it is too technical.

One of my favorite stories is when I spent days and days organizing some data in one evaluation...in tables, color coding, thinking that these lay-people, smart lay-people..I was trying to stay as close to
the data as possible so that my ways of thinking about it wouldn't be represented in the data. Well, they were just completely overwhelmed...it was as if 'what are you doing to us? What do these say?' I mean it was a real lesson. I had spent a very long time trying to keep my interpretive frame out and realized I had done that too much.T04

I wanted them to help draft the final case study but it was too time intensive for them. The data collection took a long time, analysis took a year, and the writing of the case study took almost a year.05

A lot of those people were very unaccustomed to even being asked the kinds of questions that we were asking them...and it took a lot of effort on our part to have that kind of discussion with many of the stakeholders because they were dumbfounded...even the funder!...we also asked those folks were there any particular things they thought we should do, that was very pointless! I mean they simply were not helpful at all, and we ultimately made decisions about data collection strategies, based on our own experiences.T07

I haven't been involved in too many evaluations where the clients of the stakeholders did a lot of collection of data. It just so happens that for a variety of reasons that did not work out...most of the things that I did were in much more compressed time frames, to imagine pulling that off would take forever.T12

However, having learned from these failures/limitations, the target evaluators discussed positive and productive involvement activities:

1. Having stakeholders design and collect data in professional review evaluation.

Particularly in professional reviews, then you should have a lot of input from the professionals who are under review.T02

2. Using a narrative summary of the data/results as a basis for discussion and feedback.
Now I will do things like a narrative summary of the results, which I acknowledge has my, you know, my filters in it...but I will say here are the key findings and then we will have the discussion.

I will circulate in a short term evaluation some interpretive reconstruction and I will say please don't read this like this is a story about you and your program --add to it, tell me where I am wrong...the goal here is for us all to come to some interpretive account together.

3. Involving stakeholders in designing (not writing) the final report, helping them contribute to case studies through focused writings and journals.

I have asked others to contribute narratives, to do focus writing kinds of approaches...like portions of journal entries.

They have also been involved in helping us think about how to write the report. They haven't been and won't be involved in writing it, but we have had a number of meetings to talk about audiences, format, and we have agreed on what I think will be an interesting kind of format for the report.

4. Acting as facilitator in an action research process to help build evaluation capacity for the clients/stakeholders.

If you want to do real evaluation that helps in understanding and plays a major role in being useful, and that is something that everybody mentions, but don't do much, then the client has to be an evaluator.

The first year they desperately needed my facilitation and help, the second year much less so, and this year they said 'It is nothing personal, but now we have integrated this into our practice' and so evaluation has become a way of life for them, it is very exciting.

Another way to represent the varying activities that
these two groups of evaluators involve their clients in
is to think in terms of two continuums, (1) who is in-
volved in preparing the evaluative statements/conclusion
in the evaluation process ranging from being purely
stakeholder/client/audience developed to purely evaluator
developed, and (2) when in the evaluation process are the
stakeholders/client/audiences most heavily involved,
ranging from the beginning (development of evaluation
questions stage) to the end (development and dissemina-
tion of findings), as depicted in Figures 5 and 6.

As mentioned earlier concerning continuums, none of
these persons lie strictly at only one point on the
continuum; depending on the individual evaluation pro-
ject, they may put more emphasis on earlier involvement
by stakeholders and have them develop all of the find-
ings, or the reverse may be done in another situation.
The placement of the individuals on these two continuums
represents the main emphasis of their statements about
how they involve stakeholders/clients/audiences in the
evaluation process, as presented in Figure 5.

Note that most of the evaluators cluster around the
end of the continuum which emphasizes stakeholder devel-
opment of evaluation questions/focus. This is not sur-
prising because it has become common practice to involve
those persons most affected by the evaluation in the
development of its focus and implementation.

From the stage where you look at clients, obviously they are a very important factor in understanding the problem, in finding out what the evaluation needs are, and all those things are not new. T01

Ideally you want to talk to the clients and find out what it is they are seeking and also to the other people in the program—so I would subscribe to some kind of stakeholder view, the people who have a big interest in the program. The evaluator should consider their interest and their views—what they think about the program. T02

But most often their role would be in thinking about the kind of questions we ought to ask and making interpretations of the data that we get. T12

I try to sit down with them and talk about what are the questions that they want answered about the particular program, regardless of what it might say in some particular program document. R01

We really rely on their help, we listen to them carefully because they give us cues and it helps us to be more in tune with the purpose of the program...the staff can be very helpful in uncovering some solid data...give them the opportunity to suggest things that they might think would be appropriate for use to ask in client-staff interviews. R03

They are involved in helping design the surveys and then also the information is always presented to them, always shown to them. R05

Some feedback on preliminary analyses, a sharing of some during preliminary findings, involve them in 'what do you think this means? Are there other analyses that you think would be important? Other things
that we should be looking at?' Certainly they contribute very strongly to the identification of questions that they want to answer.R06

I always try to involve whoever the targets are in terms of the services being offered...if it is not the kind of information they want, they do not pay any attention to it. So it is very important to involve them from the beginning and say 'what do you want to know and what kind of information is important?' And so I try very hard to find out at the beginning not only what it is people want to know about the program, but what kind of information will they listen to, who do they want to hear...whose voices are they interested in? I try to involve the audiences in the presentation of information.R07

--- Figure 6. Point of Emphasis of Stakeholder Involvement. ---

The differences between the target group and random group in terms of when stakeholders are involved in the evaluation process are minor, as presented in Figure 6. Both random and target evaluators focus on the beginning ("We really rely on their help, we listen to them carefully because they give us cues and it helps us to be more in tune with the purpose of the program"R03), or on the end ("I like that process of getting the critical people together to literally process the data if you

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
The only difference between the groups is that only members of the target group emphasized constant involvement in the evaluation process.

But if you talk about the total picture it is constant, there is no end to it. It just goes from the very beginning to the very end, it is just constant working with them, negotiating and understanding, there is no beginning and no end, really.

The Communicating and Reporting of Evaluation Findings

In addition to describing how they involve clients/stakeholders, the participants addressed how they communicate evaluation information and their perceptions of the importance of feedback/communications, the individuals involved, the methods used, and their satisfaction with these methods.

Each evaluator in both groups stressed the importance of face-to-face meetings, informal conversations, and creating a context for dialogue so that clients/stakeholders can take ownership of the evaluation process and the findings. This ownership is vital in building the capacity to conduct evaluations and institutionalize the process.
And then what I did, rather than write a report, was I put together the information in sort of a summary form to address the questions, and then I called all of these people together and I called it a retreat.R01

I made an oral report to all of the collaborating agencies. These are social, non-educational agencies, along with the school people and they were very receptive.R02

I try to get people to let me do it both orally and in written form. I also try to get them information before a final report...to get people to hear the information before they read it in a report and hear it in incremental ways.R07

I usually always have some kind of face to face meeting with the people...face to face meetings are important because people often misunderstand what is in the report...I've met with significant committee members in education at that time, face to face, over dinner and conveyed the essence of the evaluation to them.T02

The stakeholders need to make some claims on the findings and make them their own. Rather than see them as the 'truth' or the facts...a process is needed by which they even change the interpretation or offer multiple ideas in that process...deprivilege it, the data are only what we interpret them to be...we need a process that helps people make claims, get some ownership.T04

So I have really come to the point of relying a great deal on sort of a dialogue, as the primary means. And it is real important to think about how you create the context for dialogue. Making sure that people don't feel left out, being inclusive, having those be productive and people feeling that something has been accomplished, but without a doubt waiting until you are done is far, far too late! One of the frustrations though about doing that is that there is a great deal of impatience for dialogue...so people really are very impatient about the rhythm that a good dialogue takes and kind of developing some tolerance for ambiguity and tangents and so on. One of the things that we do hope though is that they will see something, see it as something useful and that they will do it themselves after we
are gone...so what we try to do is to have meetings where the dialogue addresses the issue that we think is at stake and then maybe they'll come to an understanding about that, if we put an issue on the table and add some information, and create a space for them to explore the ideas with us.

We set up a network of teams that met probably three times a year, where they meet and talk together about progress and problems and that is really important because it is not me facilitating at all. So there was that piece of sharing. They had to write papers, we learned that very clearly, that you have to write it down. If you don't write it down you lose it...when they have a thought about their action research they jot it on a sticky note and stick it on a pillar and when they have time together they pull down all of the sticky notes...as problems come up they are resolved quickly, whereas for anecdotal evidence, it is that too.

I probably use more often memo or draft reports...because I find in the educational field or human services it is really hard to get people together as much...so one of the things that I will do for example is particularly at the end I have a 'wrap session'. I will try to get them together and really come out with my findings and I try to locate that interpretive reconstruction of a program...and say 'O.K., let's talk about these things, what rings true what doesn't?' and so you know it continues to give them a chance as these things emerge and are reconstructed to say 'no that is beginning not to feel right'. I don't understand why evaluators are so afraid to be theoretical because I get the kind of comments back 'that I knew this but I didn't know how to express it' just in the sense that 'this really helped me to feel good, it felt right, that I have a better understanding now and a basis to begin thinking about this.' It is really just a neat experience and I can get pretty theoretical on my reports and into things that they never anticipated, but I have yet to have anyone say that they don't think that it is practical.

Some evaluators use interim memos and draft reports to communicate, focusing on the perspective and language of the readers, while other have little confidence in
written reports, using them only as historical documents.

We typically have the paper trail, so there is always typically a paper report. I always do some type of an oral briefing. Even if it is short, to walk through certain things...maybe electronic sharing, we are big on E-Mail, our state is connected by electronic mail.R01

Reports, meetings, one-on-ones when the administrators and faculty members go out it is one-on-one. They meet with all of the teachers and they go over the reports...we present at the faculty workshops...and a mail-out report is given to them.R05

I was trained as a fieldworker and have always felt that I write very simply, directly, clearly...but I am becoming more sensitive all the time to the way in which one person's simple and clear is another person's jargon!...people tend not to pay much attention until it is really getting close to something they consider down to the wire, so I could submit lost of interim reports, but they wouldn't respond to them.R06

I tend to use more informal things, fewer written things in the process. I will use more sitting down and talking with or talking on the phone with, rather than putting it on paper...I don't want to write something three months into an evaluation and them have it turn out to be totally different because I have gotten a lot of other data in the interim. So I tend to say 'Here are things that I am finding out' in an informal phone call or sitting down and chatting with people saying 'Do you think that this is verifiable, where else should I go?'R07

You definitely need informal oral interaction; you definitely need a formal report...write it from the perspective of the reader, not from the perspective of the writer...when you watch how clients/practitioners read things, they usually go to the end. So if they go to the end why not start with trying to figure out how they read it. What kinds of things do they read at home...it is not only just changing the order, that is a gimmick, but that there is also substance there.T01

The final report is a real nothing report!T05

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
I have not had a lot of confidence in written reports. I mean they serve an important kind of historical documentation need and I wouldn't berate that at all. But as means to communicate I think that they are fairly ineffectual...They have been involved in helping us think about how to write a report...so essentially we have decided to have a report that, would look like a typical evaluation report and have descriptive and evaluative conclusions, but we are also creating a kind of 'meta-text' that will be intended to be provocative, so you on the one hand have the stuff that sort of chronicles what has happened and then a text that will move people forward...we need to create a way to have some dialogue about the issues that are really relevant in the bigger picture.

So in that sense the writing of the report was not a total waste of time. But the real changes were in the lives of the people who were in the project, they had already occurred.

They are more like stories. I try to write in a way that conveys a set of themes about the project of program that I was looking at...in general I try to write careful, interesting description of things and then weave into that my interpretations of that. But again, those interpretations are grounded in the data that I have gathered and made sense of from the people that I have talked with.

A few of the target evaluators stressed, that along with the importance of creating a context for dialogue, one must become actively involved (through participation in the program/project itself, extended contact with stakeholders, and open discussion of ideas/values) in the school, project, or organization in order to build up the levels of trust, comfort, and patience necessary for productive dialogue about substantive issues.

Videos, I use slides and projective techniques to get responses, focus groups. Mostly really emersion. The fact that I am in the community or in the pro-
gram for long periods of time...that gives you patterns of behavior to start identifying over time.T03

I do a lot of site visits with the participants and so it is conversations in the hallways and conversations over playground duty...but I guess is what it comes down to is that we have to show up a lot (that is be a part of the school culture) before we can actually do any data collection.T05

One of the frustrations though about doing that is that there is a great deal of impatience for dialogue...so people really are very impatient about the rhythm that a good dialogue takes and kind of developing some tolerance for ambiguity and tangents and so on. One of the things that we do hope though is that they will see something, see it as something useful and that they will do it themselves after we are gone...so what we try to do is to have meetings where the dialogue addresses the issue that we think is at stake and then maybe they'll come to an understanding about that, if we put an issue on the table and add some information, and create a space for them to explore the ideas with us.T07

Most of the participants stated that they probably do not spend enough time on communicating efforts and perhaps should take at least as much time, if not more, for evaluation communicating/reporting as is spent on collecting data.

I guess that I could probably do a better job. I could probably be more inclusive sometimes...perhaps I should be more pro-active from the beginning.R01

Actually we don't know enough about the reporting of findings. We have methodologies about other aspects of evaluation, like data collection, data analysis. We have very little knowledge and validated instruments that communicate evaluation findings.T01

By communicating and reporting--you don't mean just writing--Well, I don't think that I devote very much time to that, probably not enough. I'd guess 5-10%. I probably should spend more time helping them work through the results.T02
Meetings are a part of the dominant culture...a lot of people don't have meetings, or go to meetings, they do it in different ways...I am a member of the dominant culture, but I recognize the problem, it is exclusionary, but it is a process that is not how a lot of people, specifically the ones you may want to include in the process, you are inviting them into your territory, you are not doing their thing.

I'm not happy with my ability to provide an executive summary or a synopsis of the work in the right bit size, in the right medium, for the people who ought to know more about it...I am not comfortable with our ability to report what we have learned.

Much of evaluation should be communicating and reporting to people...people think that data analysis is the focus, it is very small. Communicating with people is the most important part, really listening, responding to concerns. I use informal meetings...I spend as much time on that as data analysis, in 18 months of project, only maybe 3 months was data collection.

I draw the analogy between thinking about evaluation reporting and writing up the results of field studies, that we ought to allow at least as much time for that as we did in collecting the data, if not more, and I think that is a critical part. I always try to build in time to talk about them before, in draft form, or them to review them and so forth, I think that that is critical.

Here again, the differences between the random and target evaluators are slight in terms of the type of communicating and reporting activities that they use and in terms of the importance that they place on communication in the evaluation process. The main difference is that only members of the target group stress the need for longer-term, active involvement in the program itself as a means of building up trust and patience for productive dialogue. This echoes the earlier discussion where mem-
bers of the target group emphasized in the model or method of inquiry that they used being planted in the community in order to build continuous long-term relationships with program participants and being involved in the program so as to experience the same issues facing the program deliverers and recipients.

Ethics and Social Responsibility

The interview also focused on the personal views/values of these evaluators and the extent to which these enter into their choice of evaluation projects and how they conduct an evaluation, and the extent to which these view/values are articulated to clients/stakeholders. The participants were also asked if the way in which they conduct evaluations poses particular ethical dilemmas, and if so, how they approach/resolve these dilemmas.

Each of the target (sentient) evaluators see evaluation as having some type of social responsibility:

1. Evaluating congressional decision-processes and policies.

I think that it should be to determine the merit or worth of something but I think that they should be conducted at all levels of society, so by that I mean evaluation should occur at the highest levels, Congressional or Presidential level as well...if we were doing it properly then the overall effect would be to have evaluation at all levels and that evaluation would be public information...I see evaluation
as actually the mechanism for informing and judging a lot of things that are going on in society. I think that this is a broader view...I think that most of them [evaluators] would tend to see themselves as doing a program, approaching the client...and they would approach that person and say 'what is it that you want me to do?'. Now I think that is totally inadequate. I mean that I am in favor of doing that, involving the client or the sponsor, as they are a major stakeholder, but that is not enough you have to think about the larger public good as well and other people who are involved.

2. Facilitating action research at the "grassroots" school building level.

It has always befuddled me when teachers say that it is not the role of the school to take on society's problems, that is society's problems! That should be the goal, the people go on and take action themselves, evaluate their own endeavors because they have come to see the value of reflection, interpretation, and investment in themselves and those around them.

I am much more now into the grassroots, programmatic change that I think over time adds up to social change.

3. Promoting and creating socially responsible institutions.

One of the main purposes of evaluation is to improve society. From my perspective it is to do some social good. So it is a social and moral responsibility, there is no question about it. It is definitely a part of evaluation...moving forward in a progressive fashion.

Absolutely, in promoting and creating socially responsible institutions, in my case, schools primarily. I think that it is the case that all ways of thinking about evaluation entail a notion of improvement. I've called it the ameliorative assumption. Any approach to evaluation presupposes that there is going to be some amelioration, that things are going to get better, that is a built in expecta-
tion. Now what that means is quite different depending on what you do...Now the degree to which you see that as a technical act or a moral imperative seems to differ in people. And I certainly see it as a moral imperative. It is not about the exercise of science or 'truth' that you are about, it is about moral obligations more than anything else.

4. Creating a public conversation about the purposes, goals, and values of programs/policies/institutions - a social critique.

To me it is just that simple, if you acknowledge that this is a process with a political stance and values, then what values and stance do you want to promote?...I was in a panel yesterday I really do think that it was a very manager, managerial point of view, and I would really reject that as a primary role for evaluation, it is ok but it shouldn't be our main role in society. But I think that is most people's view.

Think about the complex ethical dilemmas of the use of Norplant...this is a medical procedure, to what extent could this become coercive, the evaluators have a responsibility to address such problems...the technical-rationalist design is not going to work.

Evaluation is ideally positioned to do this and therefore it ought to facilitate this sort of public conversation about the purposes and goals, and the value of what we are doing...I see the kind of evaluation that I am talking about to be closer to sort of social science's public philosophy...I think that evaluators ought to look more like people like that and step up the public discussion about the purposes and the means and the ends, and goals of programs...I think that most people still think of evaluators as technical experts or social scientists who gather data to answer questions or help decision makers make decisions...I think that the idea of gathering evaluation information in order to have a richer conversation is not something that most of my evaluator colleagues think of as a worthwhile purpose of evaluation. I just want to add that I think that the relatively newer generation of evaluators does think like that.
The random evaluators varied in whether they saw evaluation as having social responsibility. Most saw evaluation's role in terms of improving the program or services offered in order to improve people's lives.

I think that there is a notion of improving that I have talked about, program improvement; I think that there is also a notion of fiscal and social accountability...there is this social accountability for saying we should be using those to better people's lives and we should be demonstrating that if you will. So the social-fiscal are sort of linked, I think it goes along with this notion of program improvement.R01

If the client group that is supposed to be benefitting and if the deliverers decide that at any time it is not worth it, then it is over, it is dead in the water...So I am concerned about how they feel...but I am certainly not in the camp of the marxists.R02

Even when I do research, pure research, in my own sense I always try to leave the setting or the people in charge something that they can work with to improve the program.R05

Only one member of the random group touched on the notion of a social critique, but again it was focused more at the programmatic level than at the societal level.

I think that evaluation has an important role because it can change people's thinking from 'how do I save the program' to 'how do we best meet the needs of a certain group that we have tried to address'. And there is a social responsibility I think for us to focus on who we are trying to serve and what it is we are trying to do and orient evaluations in that direction, more than it is to say this program is good or bad. It is also very hard to do because the programs pay you to do the evaluation.R07

This discussion reflects back to Figure 4 which
shows only members of the target "sentient" group as having a macro-global-societal focus for the purposes of evaluation. This may be putting too fine a point on it, but the random group's focus tends to cluster around the micro-program level, while the target evaluators also sense a larger role and responsibility for the profession.

Most of the evaluators have flexibility in choosing their evaluation projects, either because they have faculty positions or other income or have developed reputations in their areas of expertise so that clients actively seek them out. Although each mentioned that they have been involved in evaluations that had conflicts (either because of personal or philosophical differences), most stated that they now only take on evaluations that interest them and that reflect their values.

Interestingly, one target evaluator stated that although s/he only chooses projects that are interesting to her/him, s/he did stress that the profession of evaluation has a responsibility to provide evaluation services to all agencies and needs to discuss how this can be accomplished, perhaps with a system not unlike "pro-bono" legal cases.

Specific ethical dilemmas and conflicts experienced by the target evaluators include being asked to change or
"soften" findings, as well as the following:

1. Wording or phrasing in evaluation reports that may be unknowingly offensive or disturbing to clients/stakeholders.

Now sometimes you find that if you show them the report and they get angry and you ask them what they are angry about, sometimes it will be a phrase or a word that carries a meaning to them that you really didn't mean it to mean. Sometimes you just change the phrasing around and you take some of the conflict out that way.

2. Personal, sensitive information collected through case study/ethnographic methods and how to protect clients/stakeholders who may unknowingly risk their own confidentiality.

In employing case studies you can find out a lot of stuff—much more than you can in other kinds of evaluations. You find deeper stuff and more personal stuff, so you always have the problem of how to use that in a report...so what should you report and what should you not report? I think that is a big problem...well there are certain kinds of principles that I would follow—like not to say something about someone's personal life that would be derogatory—unless it would be in such a way that strikes at the actual center of the program or policy under review—the rule of thumb there is don't question people's motives.

I think that the way that I do evaluation tends to get into the personal affairs of teachers, administrators, and other people. A lot of probing, a lot of informal data gathering that does put people at risk and exposes them, sometimes in ways that they don't really seem to mind being exposed, so we cannot count on them being protective of the confidentiality that they should have. That is an ethical problem that I and my colleagues sometimes face.

The nature of qualitative data when you have a primary informant that doesn't represent numbers, but
represents weight and it is really critical that people understand who it was that said this, not in a personal sense, but in a job-type area or role kind of way. That by hiding or obscuring it in the report so that confidentiality isn't lost, it just looses all of its punch! I try to think of ways to get around that on the front end, so that I get triangulation right away.

3. Discussing participatory issues/methods without coming on "too strong", how to introduce a wholly different paradigm, how to discuss out-dated notions of evaluation, how to discuss our values systems and bring in literature to frame the discussion, and how to re-educate sponsors and clients.

I said maybe we can look at some of the participatory issues that I am interested in...and I made it really clear, well I have never really done this so clearly before, but I thought that my interest in participatory values probably needs to be said. I'm not sure I would do it quite the same way next time, but I did say that one of our agendas would be to try out these kinds of processes...So I don't think that little talk was well received there...I think that I kind of over did it...I think that it is important to do so, but probably not quite so heavy-handedly. I think that the way it was delivered was inconsistent with the message.

Some projects at the university were supporting a very narrow construction of learning mathematics for students...we were flying out there and doing this and leaving and to me it seemed very positivistic. I didn't feel like a decent person to do that...it is artificial to me. It is simply not legitimate...we are talking about very different paradigms...it is a wholly different paradigm and you need to situate yourself in it and look from outside of it.

So I inherited some designs and some expectations about evaluation that were hard to live with...what I did ultimately was to adopt a comparative framework, do away with some notions of experimentalist, change some of the language.
They had some pretty specific ideas about what they thought evaluation should be...the concept of working with people with out-dated notions of evaluation is important...it took a lot of education...I said what I think we are doing is important and if we have different ideas we need to talk this out. I see a tension there in talking about being an advocate, how do you be an evaluator and go into a project and talk about what your value systems are, and should you?...so I would like to see more examples, explicit recognition of value systems by evaluators, what that does to the evaluation process.T08

How far do you go in injecting your viewpoint? How do you frame the debate in ways that they did not see? You have to bring in literature. How much of my view is white middle class structure?T11

4. Providing timely findings, while trying to keep clients/stakeholders from viewing interim memos as "final" statements.

I used kind of what I called 'reporting memos' that I would write to project directors on an ongoing basis...but ultimately there were several instances where project directors would take those memos and use them as if they were sort of final statements...so I always felt very uncertain about waiting until the end...I tend to rely a lot more on face-to-face dialogue.T07

5. Validating your reconstruction of data into a framework, is it really validated by the clients/stakeholders? or does it simply sound logical to them, thus they accept it.

In my reconstruction, because I make leaps of inference sometimes, sometimes I'm feeling in a lot of gaps there, based on what I see human agency is meaning in this situation. I am assuming that the reconstruction and the context, and the phenomenon is this, and I'm trying to validate that with the stakeholders. I'm not trying to make a direct appeal to evidence, though I'm using evidence to shore up this thing, that doesn't shore up your interpreta-
tion, that only shores up the facts that are in your interpretation...and I go 'gee have I pulled one over on anyone here?...'...because it makes so much sense? I mean I can make anything look like it makes sense...I could be dead wrong about the interpretation...but because sometimes it makes so much sense they will be very excited about it and I'm not always sure that that is the same as validation.

6. Cultural differences in educational policy and theory.

Which is what was believed in America at the turn of the century!...so we clashed on the idea of persons with disabilities being capable of learning.

If for example you were looking at questions of discipline and behavior, one of the real challenges in the building is how do you accept the African-American culture, but at the same time demand academic behavior so that they can learn, and that is where I can see getting into ethical discussion about whose values are we going to base our assessment of behavior on?

Specific ethical dilemmas and conflicts experienced by the random evaluators include being asked to change or "soften" data ("But what I ended up doing was going back and I did change my report somewhat, and I softened it, and after I did that I just felt like a prostitute!" R01; "We took it to the State Legislature and they said 'we don't like these. Turn the data around.'" R04) as well as the following:

1. Personal dislike for a person involved in the evaluation process or working with someone who has questionable personal motives.

I worked with a person who just bugged me! I think that I conducted myself very professionally, but
this person just drove me crazy!R01

The program administrator, a white male, who seemed to be used to telling people what to do...we walked in and to our surprise he told us how the evaluation was going to be done and what instruments we would be using...and we had to kind of say 'excuse me, I think we need to talk about this because we were hired to do this and we are going to have to be the people who make these kinds of decisions and this is our design and if there is something that doesn't ring true with you then we will talk about it, but we have to make the choices about how we collect data, what kind of data we collect.' I think that some of it is maybe resistance on the part of the particular individual, frequently because they have had a bad experience prior to that, this evaluation.R03

2. Being asked to do an evaluation in order to "validate" the program.

Now I have encountered some that I suspected strongly...I knew that their view of evaluation was 'well we already know that it is a good program, all that we want you to do is get the evidence so that we can show that this is a good program.R02

They may say that I hope you will present us in a good light, and I will say something to the effect that 'I can't say more or less than what the data say, we can talk about language that I use, and I will tell you the limits beyond which I will not go...there are definitely pressure, but I have to draw very clear lines.R06

I won't take on evaluation jobs where I perceive it is being done to meet a letter of the law, rather than the spirit of the law. To some degree I am not interested in being a functionary, doing a functionary task. R07

3. Dealing with existing data and reporting about it in politically charged areas.

In criminal justice research you use a lot of existing data, and I don't consider that to be the strongest data cited, but it is the best I am going to
get to do the project. On one occasion, because some of the work that I do is very highly politically charged...the media began to pay a lot of attention and began to distort things to some degree.R06

Even in oral presentation I am very careful about it and I try to do them in a way that isn't very pub­lic. A newspaper article totally misinterpreted what we said.R07

4. Using personal, sensitive information.

I found out some data that was purely confidential that a teacher was really inadequate...I felt like somebody really needed to know this, nobody seemed to be gathering the information...on the other hand the data had been gathered in a purely confidential manner. The way that I resolved it was to find another way to get the same data...but I still never really felt comfortable in that I used the information...It is sort of like I am never really supposed to ever use this because it is strictly confidential in the way that I had gotten it.R07

The ethical concerns and social responsibility men­tioned by these two groups, can be represented in the following continuums presented in Figures 7, 8, 9, and 10. The differences are most striking in Figure 7 in terms of focusing on ethical dilemmas concerning the responsibility of reeducating clients and sponsors and in Figure 8 in introducing a discussion of participatory evaluation and the values/philosophical stances.

---

13 Note that not all of the participants of the study appear on the continuums presented in Figures 7-10. Only those members of the random and target groups who made specific, direct statements related to the topics presented in the figures are included on the continuums.
<table>
<thead>
<tr>
<th>Target Group</th>
<th>Random Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T02, T03, T04, T05</td>
<td>(R01, R02, R03)</td>
</tr>
<tr>
<td>T07, T09, T11, T12</td>
<td>R07</td>
</tr>
</tbody>
</table>

Figure 7. Social Responsibility of Evaluation in Terms of Reeducating and Discussing/Introducing Notions of Participation.

<table>
<thead>
<tr>
<th>High</th>
<th>Low</th>
</tr>
</thead>
</table>

Target Group
(T04, T05, T07)
T08, T09, T11

Random Group
(R05)

Figure 8. Need to Have Theoretical/Grounded Pragmatic Dialogue or Discussion of Values.

<table>
<thead>
<tr>
<th>High</th>
<th>Low</th>
</tr>
</thead>
</table>

Target Group + Random Group
(T01, T02, T11) (R01, R04, R06)

Figure 9. Resistance to Changing Evaluation Findings/Validation.

<table>
<thead>
<tr>
<th>High</th>
<th>Low</th>
</tr>
</thead>
</table>

Target Group
(T02, T06, T10)

Random Group
(R05, R07)

Figure 10. Sensitivity to Using Personal Information.

Future Issues for Evaluation

Finally, the participants were asked what they considered to be the most salient issues for the future of evaluation and how those issues might be addressed.
Five main areas of concern emerged from the target group:

1. Focusing on underrepresented groups, minorities, the powerless, and the suppressed voices and how to balance the diverse voices;

2. Discussing the substance of what evaluation is or should be, theoretical discussions of who we are and what we do, discussing our values through critical analysis;

3. Institutionalizing evaluation in programs/organizations, becoming a part of people's thinking, planning, and management on a regular basis;

4. How to be appropriately "negative" -- tend to be too positive -- need to be able to say when something is bad without developing an adversarial relationship -- Catch-22 situation -- re-educating sponsors and clients; and

5. Evaluation as a reform strategy, the need to address mainstream social concerns and understand history of programs/policies, and discuss political/activist nature of evaluation.

For the random group three areas of concern emerged:

1. Institutionalizing evaluation in programs/organizations, becoming a part of people's thinking, planning, and management on a regular basis;

2. Focusing less on the qualitative/quantitative arguments and more on more effective ways of measuring
interventions; and

3. Being more responsible to and empowering of clients/recipient/"learners", including looking at cultural diversity.

These future issues discussed by both groups serve as a springboard to thinking about the implications for the evaluation profession and educational evaluation.
CHAPTER V

CONCLUSIONS

Implications

The differences between the target groups of sentient evaluators and the random group of evaluators were at many times very subtle. But it is the subtle nuances that separate the sentient evaluators from other evaluators which may distinguish the sentient worldview/framework.

The differences were slight in terms of (a) standard evaluation practices (involving stakeholders/clients/audiences, establishing communication mechanisms, and making evaluative statements of merit or worth); (b) the joint development of the evaluation questions by the evaluator and the stakeholders/clients/audiences; (c) the importance of meetings, informal conversations, and creating ownership of the findings in the evaluation process; (d) not spending enough time/effort on communicating and reporting activities; (e) resistance in altering evaluation findings; and (f) sensitivity to using personal information in evaluation reports.

Where the two groups, the target group (evaluators defined as sentient) and the random group (chosen from
AEA membership), seemed to diverge most strikingly is in (a) seeing the purpose of evaluation as having a global-societal responsibility; (b) noting the need for evaluation to create a social critique; (c) promoting the constant involvement of stakeholders in the evaluation process; (d) encouraging active, long-term involvement in the project/organization to build up trust and commitment; and (e) seeing the need to have theoretical discussions/grounded pragmatic dialogues of values and how values affect the evaluation process and how they are communicated to stakeholders/clients/audiences.

These differences suggest the following hypotheses that could be utilized in future sample-survey studies from which generalizable inferences could be made:

Hypothesis 1: Sentient evaluators differ from other evaluators in terms the context they ascribe to evaluation (macro-level or micro-level of focus);

Hypothesis 2: Sentient evaluators differ from other evaluators in terms of evaluation's role in promoting social criticism;

Hypothesis 3: Sentient evaluators differ from other evaluators in terms of the involvement of stakeholders in the evaluation process;

Hypothesis 4: Sentient evaluators differ from other evaluators in terms of the length, breadth, and depth of
involvement in the project/organization being evaluated; and

Hypothesis 5: Sentient evaluators differ from other evaluators in terms of evaluation's need to have theoretical discussions of values (how they affect the evaluation process and how they are communicated to stakeholders/client/audiences).

One main weakness of this study is that it is based on only the "word" of the participants in both groups as to whether they actually "practice what they preach," because the researcher did not interview any of the clients/stakeholders/audiences that they have worked with, nor were existing evaluation reports, videos, memos, minutes of meetings, or other types of documentation of evaluation activities and communicating/reporting strategies collected. The collection of such documents and the contacting/interviewing the stakeholders would be a complex and exhausting task, both legally and logistically, certainly not impossible, but considered to be beyond the scope of this study by the researcher.

Thus, one might argue that the differences between the two groups may be more a matter of language/vocabulary and comfort/experience in discussing the issues raised in the interview protocols, than a difference in actual practice. However, there is some evidence
in their writings/presentations (part of the criteria/sources that were used to choose the target group for possible participation) that these persons have a different viewpoint/perspective on what evaluation is and what it should achieve. Also, all participants were able to change or make additions to their transcripts at their leisure and thus could have reflected on their comments in a more comfortable setting, choosing whatever language/vocabulary they saw fit. None of the additions/changes made by the participants on the transcripts radically changed the information/issues covered (most were concerned with typographical errors or cleaning up gaps in natural speech patterns, i.e., "ums," "ahs," and fragments/run-on sentences), and therefore this study assumes that all participants were comfortable with the way that their ideas and activities were portrayed in the transcripts and, as much as a one-time interview can, presented an accurate portrayal of their evaluation practice.

Another weakness of the study may be that one could argue that sentient evaluation is the luxury of the academy in that faculty members do not have to rely on evaluation contracts for their livelihood and as a result can create or choose evaluation contracts that may be more unusual or theoretical in nature. In fact two of the
targets suggested this possibility:

I have the luxury of the academy. It is very easy for me to sit in my little office and teach social justice oriented evaluation...in my defense I also try to do it...but practicing evaluators have real multiple constraints and demands, and the reality of the rent to pay.T04;

I have the luxury to do that because it is not my main role. I probably would think differently about that if that was the way I made my living.T12.

Although the researcher does not argue with the notion that the financial considerations are different between a full-time internal or external evaluator and a faculty member/university staff person who does evaluation as a side-activity, it is interesting that the one member of the target group who earns 100% of her/his income as an evaluation consultant directing her/his own evaluation firm made the following statement:

I don't understand why evaluators are so afraid to be theoretical because I get the kind of comments back 'that I knew this but I didn't know how to express it' just in the sense that 'this really helped me to feel good, it felt right, that I have a better understanding now and a basis to begin thinking about this.' It is really just a neat experience and I can get pretty theoretical on my reports and into things that they never anticipated, but I have yet to have anyone say that they don't think that it is practical.T10

Looking at the broader implications of the study for the field of evaluation, none of the evaluators who participated in the study were completely content for evaluation to follow along a rigid, techno-rational, evaluator-as-expert road. These evaluators saw the need
to address cultural diversity and find ways of having evaluation institutionalized in organizations and programs and processes, instead of as an "add on".

If education (focusing on this arena because most of the persons interviewed in this study practice educational evaluation) continues along the course of reform and systemic change, "one-shot, outcomes-oriented, expert-objective" evaluation practice simply will not work. Because the process of systemic change is slow, complex, non-linear, and context-linked, a long-term, indepth relationship needs to be established between an evaluator and school/program personnel.

The role of the evaluator becomes more of that of a sentient who assists the organization (district, school, program) in developing an "ethic" of continuous reflection and questioning in order to improve the life of their constituents and examine their contribution to society, rather than as an expert-technician who has a limited, prescriptive relationship, based on short-term involvement.

The responsibilities and skills of the evaluator slowly become integrated into the organizational system, instead of being "housed" in one individual or small evaluation team. As the stakeholders and evaluator jointly reflect on the focus, needs, and future of the
organization, by revealing individual beliefs, needs, and skills, the role of the evaluator can be transferred to the organization, as it grows more comfortable and skillful in understanding itself and its context.

This focus on the institutionalization of an evaluation ethic, necessitates alternative communicating and reporting strategies; relying on a final, comprehensive technical report is too little, too late. Communication takes place throughout the evaluation process, whether through face-to-face contacts, informal conversations, memos and draft/interim reports, or forums, so that feedback can be used for continuous improvement and elucidation.

Relying on the facade of objectivity and the technical-expert persona will not promote the discussion of the profession's most critical issues. As opposed to the "have-gun-will-travel" movement from contract to contract, evaluators would become knowledgeable about the context/content of the evaluands and their historical and socio-political context in order to promote quality and accountability by providing informed recommendations.

The implications of this study's findings for educational evaluation include the need to (a) question purely summative evaluation conducted by short-term external entities; (b) establish long-term, indepth relationships.
between evaluators and school/program personnel; (c) create a context for continuous communication, feedback, and dialogue; (d) develop internal capacity for designing, conducting, interpreting, and utilizing evaluation; (e) address underrepresented groups and minorities and recognize suppressed voices; and (f) discuss the substance of what evaluation is or should be and the values that underlie programs and policies.

Formal evaluation has not become a day-to-day, institutionalized part of education (it could be argued that informal evaluation—determining merit or worth—happens continuously in any organization). It has not vigorously discussed its political and value stances. After 10-15 years, the battle has been won to include diverse stakeholder views (Scriven, 1993), but we have no explicit theories on how to balance these voices. The movement needs to be to balance the technicalities of evaluation with questions such as "what is evaluation? what will be the role(s) of the professional evaluator? what is good and just and right?" The worldview of the sentient as presented in this study may serve as a building block to help the field begin to discuss what such "movement" might look like, what the implications will be in terms of how evaluators are trained, and what other fields/professions evaluation looks to for ideas-methods.
Sentient

This final portrait of the sentient evaluator is created by using information gathered from the literature, the criteria/resources used in choosing the target group, and findings that emerged from the data.

Essentially a sentient evaluator takes an active role in counselling the program that s/he is investigating, deliberately seeking to represent underrepresented stakeholders or audiences. This type of evaluator has her/his own ethical agenda through which s/he views the program being evaluated and tends to have a formative perspective on program development and improvement.

The sentient would agree that evaluation should promote excellence in educational programs and that each person's essential dignity must be respected. However, the sentient might take a more directed approach to who is being served and how they are being served by specifically looking at underrepresented groups or marginalized viewpoints.

Sentient evaluators to some extent assume that the world is complex and that programs are embedded in social and political processes and organizations. Participants are involved in designing, conducting, and disseminating the evaluation. These actions are common in the evaluation profession including using multiple stakeholder
views and directly involving clients in designing and conducting the evaluation; describing the technical, social, political, organizational, and economic context of the program using multiple sources; and providing clients and other key audiences with periodic reports on the progress of the evaluation, in order to enhance the credibility of the evaluation.

While challenging the social, political, personal structures and beliefs in the evaluation context, the evaluator, identified as sentient, also simultaneously challenges her/his own structures and beliefs. S/he simultaneously reveals her/his own agendas, biases, beliefs, and understandings as s/he reveals those of the context. The sentient takes the notion of values identification further than just focusing on different values used in interpreting the data; the sentient pushes the identification of values to a deeper discussion of what values are dominant in the evaluation context and how these values play-out in the society. Closely aligned to cultural feminism, which emphasizes a holistic and collective orientation to world and work experiences, the sentient creates a kind of praxis, with continuous cycles of critique, knowledge construction, and social action.

The sentient evaluator may be reflected in some of the concerns behind creating "morally engaged" evalua-
tion. This type of evaluator might agree that human inquiry cannot be divorced from historical, moral, and political locations and that politics and science are intertwined in evaluation practice. The sentient evaluator may also agree that those who conduct research within an evaluative context are obligated to consider the moral implications that their findings may carry for those who are affected by the program. These evaluators acknowledge to their audience their position on issues of social justice by engaging in self-conscious reflection of their values and articulating these values.

Sentient evaluators are self-critical, identify the moral issues and work to resolve moral problems in their own work, and behave in morally responsible ways. They might say that being ethical in program evaluation is a process of growth in understanding, perception, and creative problem-solving ability that respects the interests of individuals and society. These evaluators help the respondents cope with their new knowledge, creating situations without deception in which the evaluator and evaluated are involved in the construction of reality.

Perhaps the best way to summarize the sentient evaluator reflected in the data collected and analyzed in this study is to highlight six characteristics or activities discussed by the 12 members of the targeted (sen-
tient) group: (1) promote advocacy for improving human life; (2) address and articulate the social responsibility of evaluation; (3) develop continuous long-term relationships with program participants so as to experience the same issues facing the program deliverers and recipients; (4) create a context for dialogue so that clients/stakeholders can take ownership of the evaluation process and the findings; (5) involve clients/stakeholders in meaningful, appropriate evaluation activities to build future capacity; and (6) assist clients/stakeholders in internalizing/institutionalizing evaluation.
Appendix A

Interview Protocol
INTERVIEW QUESTIONS

Introduction: Thank you for agreeing to participate in my study. Just to remind you about the purpose of my study, for my dissertation I am interviewing a group of evaluators that I have named "sentient" in order to better understand their role in the profession of evaluation. I will be taping our conversation today. Only I will have access to the tapes and only I will complete the transcription and coding processes. I will be sending you copies of the transcript and coded data for feedback and revision. Do you have any questions or concerns before we begin?

Question Area 1: Let me first ask you a few questions about your background in evaluation:
- In what settings/professions/areas have you conducted evaluations?
- In what area(s) do you most frequently work?
- In what area(s) do you plan to work in the future?
- For how many years have you been working in the field of evaluation?
- In what other professions do you work? - i.e.; higher ed., government, etc. What do you consider as your main professional role?

Q A 2: Could you describe a typical evaluation process that you frequently use?
- Do you have a particular model of evaluation or method of inquiry that you most frequently use?
- Do you communicate this to your clients/audiences/stakeholders? When? How?
- What do you see as the main purpose of evaluation?
- How does this view reflect your values and background?
- What do you see as the role of the profession of evaluation in terms of society/social responsibility?
- Do you think that this differs from most of your peers/colleagues?

Q A 3: What role do clients play in your evaluation process?
- audiences?
- stakeholders?
- Are clients/audiences/stakeholders involved in the development of the evaluation plan? How?
- in the collection of data? ...the analysis of data?
- in the presentation of findings (writing reports/presentations)
- in dissemination?

Q A 4: How much importance do you place on communicating and reporting efforts?
- How much of your total evaluation time to you spend on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
com/rep efforts?
- Who are the individuals involved?
- What kind of com/rep activities do you use?
- How satisfied are you with these activities?
- What feedback have you gotten from clients, audiences, stakeholders concerning these activities?

Q A 5: Earlier you said ....... was the main purpose of evaluation. Does this view/value, enter into your choice of evaluation projects?
- Have you ever turned down a client because their evaluation purposes conflicted with your views? Did you try to resolve these conflicts? How?
- To what extent do your views/values enter into the evaluation process?
- Do you articulate them to your clients/audiences/stakeholders? How?

Q A 6: Do you think that the way that you conduct evaluations or your view on the purpose of evaluation poses any particular ethical dilemmas? Could you give me an example?
- Do you have any particular methods that you use to help resolve or soften these dilemmas?

Q A 7: What issues do you see as most salient for the future of evaluation?
- Do you think that your views/methods/activities in evaluation can help focus/expand these issues?
Appendix B

Cover Letter and Overview
I am a doctoral student at Western Michigan University in the Department of Educational Leadership's Measurement, Research, and Evaluation Program, and I currently work for the evaluation team Science and Mathematics Program Improvement at Western Michigan University. To complete my Ph.D., I am currently working on my dissertation research which focuses on the roles of evaluators and the unique dilemmas these roles pose in how they involve clients, stakeholders, and audiences in the evaluation process and how they communicate their findings, especially in terms of the language they use and the "voices" (perspectives) they attempt to capture.

I would like to invite you to participate in my study. I have attempted to identify a select group of evaluators through their professional publications; presentations at AEA and AERA and their interest groups involving qualitative research and methods, evaluation use, and research on evaluation; and recommendations by practicing evaluators.

I will be conducting one-to-one interview sessions (~45-90 minutes in length) at AEA's annual meeting in Dallas, TX November 3-6, 1993 at the convenience of participants. As needed follow-up interviews/discussions will take place via phone calls. If you do not plan on attending AEA, alternative arrangements can be made.

I hope that this research process can be interactive with an open exchange of ideas and issues. I will be providing all participants with copies of transcripts, coding schemes, draft analysis, and written descriptions for review and revision.

I have included a brief abstract outlining the study, its purposes, procedures, and schedule. I will contact you within one week to further discuss the study, answer any questions you might have, and ask for your participation.

I can be reached at the above phone number evenings and weekends and at 616-387-3791 daytimes if you wish to speak with me sooner.

Thank you for your time and consideration.

Mary E. Piontek
Dissertation Research Overview

Purpose

The overall purpose of this study is to look at the ways evaluators involve stakeholders, clients, and audiences in the evaluation process and communicate their findings to these various groups. I will focus on evaluators who consider themselves "sentient", as well as other evaluation roles, and the ethical dilemmas that these roles pose in (a) how they involve their stakeholders, clients, and audiences in the evaluation process (context) and (b) how they communicate, especially in terms of the language they use and the "voices" (perspectives) they attempt to capture (format/content).

Evaluators are just beginning to explore evaluation that is based on anthropological and feminist theories about research, its purpose, and its "authors" or "voices". Essentially a sentient evaluator takes an active role in promoting the program that she is investigating and deliberately seeks to represent underrepresented stakeholders or audiences. This type of evaluator has her own ethical or moral agenda through which she views the program being evaluated and takes a formative perspective on program development and improvement.

My intent is neither to promote or criticize this role, but rather to develop detailed descriptions about the role in terms of communicating and reporting efforts, the involvement of various groups in the evaluation process, and ethical considerations, in order to better understand its implications for the profession of evaluation.

Little has been written about the role of the sentient evaluator, and even less discussion has focused specifically on the ethical situations encountered by these evaluators in the evaluation process. This study will contribute to this body of knowledge, as well as address some of the more salient issues of discussion in evaluation today: collaboration as a means of involving audiences, clients, and stakeholders to increase the relevance and utilization of evaluation processes and findings; alternative means of communicating and reporting evaluation findings; and the development of ethical standards for the profession of evaluation.

This study builds from the findings of a Spring, 1992 study of a random sample of 343 members of the American Evaluation Association (Torres, Preskill, & Piontek, 1992) in which evaluators described aspects of format, content, and process which contributed to successful communicating and reporting: (a) formats providing for easy/quick assimilation of information; (b) content addressing both positive and negative findings, qualitative and quantitative aspects of the program, and specific recommendations for change; and (c) processes involving clients/audiences/stakeholders in the evaluation from its beginning.

This study will investigate how the role of the sentient evaluator impacts a) the aspects of format, content, and process which contribute to successful communicating and reporting and b) the use of collaboration and the involvement of clients, audiences, and stakeholders.

Methodology
Subjects

This study will utilize in-depth semi-structured and unstructured ethnographic interviewing techniques on a select population of evaluators identified by 1) their own statements and/or published research, or 2) by their involvement in/presentations for the Qualitative Research, Women in Education, and Research on Evaluation Special Interest Groups of AERA, or 3) by their involvement in and presentations at American Evaluation Association Topical Interest Groups of Utilizing Evaluation and Qualitative Methods, or 4) by recommendations by practicing evaluators.

Where possible documents will be collected from each interviewee concerning evaluation findings, communicating and reporting efforts, and meetings or conversations with audiences, clients, and stakeholders to get additional data sources to describe, define, and clarify the final description of the sentient role and communicating and reporting efforts. The evaluators will be asked to remove all identifying information from the documents in terms of the name of the company, client, or funders; the type of business/educational/government/social service practice; and personal identifiers such as names or titles. These documents might take the form of professional presentations, papers, or articles; evaluation reports; and descriptions of evaluation contracts, contacts, meetings or evaluation designs/action plans.

The interviews will be a combination of three types of interview instrumentation: informal conversational, interview guided, and standardized open-ended. The development, piloting, and final revision of the interview protocols involved using the information gathered through literature review, past evaluation work and research, and feedback from practicing evaluations concerning the accuracy, relevancy, and appropriateness of the language of and issues in the questions. The following areas of inquiry are the major focus:

* model of evaluation used/method of inquiry or investigation
* role of stakeholders/clients/audiences in evaluation phases (level of involvement in dev. of evaluation plan, data collection and analysis, writing of "reports", and communicating and reporting findings)
* communicating and reporting findings - perception of importance, individuals involved, methods used, and satisfaction with methods
* ethics and social responsibility - personal views/values, extent to which these enter into choice of evaluation projects, extent to which these enter into the evaluation process, extent to which they are articulated to audiences/clients/stakeholders, etc.

The researcher will use a portable tape-recorder with microphone to tape the interviews within a closed setting (private room or office). Only the researcher and the interviewee will be present during the interview session which will be approximately 45-90 minutes in length. All tapes produced during the interviews will be the property of the researcher and will be only in her possession before/during/after the interview sessions and research process. The tapes will be identified by the date of the interview, the setting, and the interview identification number (ID#). The identification number will be known only to the researcher and the interviewee so that transcripts and interpretive descriptions can be given to the individual interviewees for feedback and revision. The code will be a three-digit number (e.g., 003) assigned to the interviewee by the researchers and secured in the researcher's
personal files so that only she can match the ID# to the interviewee's name.

Although all efforts will be made to assure the confidentiality of the information, it should be noted that it is not anticipated that the information collected during this research project will be "controversial or sensitive" or in any way will jeopardize the well-being, livelihood, or professional standing of the participants. As a result, the interviewees may choose to be identified by their names. If this is their preference, I will do so in my final data presentation.

Data Analysis

The interview transcripts, and any other documents/written materials, will be analyzed using an emergent design, in which a series of reflective and evolving data analysis phases will take place in order to capture the varying themes, categories, and/or "stories" that emerge from the data. The analysis of the interview transcripts and other documents/written materials will follow the same basic design presented below in Figure 1.

FIGURE 1
DATA ANALYSIS PROCESS

[Document --] or [Interview --> Transcription --> feedback from participants]
 Preliminary reading and identification of themes --> Development of preliminary coding scheme --> Use of coding scheme on transcripts or documents --> Analysis of coded data to identify themes or categories --> Review and revision of coding scheme --> Use of coding scheme on transcripts or documents --> Analysis of coded data --> Development of categories and themes --> Development or writing or "stories" or descriptions --> Distribution of "stories"/descriptions to interviewees for 'member checks' --> may lead to Second round of interview/followup --> Use of coding --> Revision of "stories" --> 'member checks' --> Final "stories" on communicating/reporting, collaboration/relationships/ involvement, ethical considerations, and other categories that might emerge.

The researcher will do all transcription of the interview tapes within a closed setting (her personal home office) so that no other individuals have access to the "raw"information. Any names or identifying information discussed on the tapes will be substituted for using the interviewee ID# or noted using dash marks or pseudonyms (e.g.) Western Michigan University = ----University). The transcriptions will be housed on computer discs accessible only to the researcher and identified only be the interviewee identification number. The transcripts will be stored in this format to facilitate the data analysis using ETHNOGRAPH (Qualis Research Associates, 198), a qualitative data analysis software package.

The researcher will review the transcripts to develop a preliminary coding scheme. Printed copies of the transcripts will be sent to the individual interviewee. These printed copies will only be accessible to the researcher (and obviously the interviewee). The purpose of providing the interviewees with copies of their transcripts and drafts of analysis description or stories is to have them review, revise, confirm, or reinterpret the information.

The researcher and the interviewee will discuss any major
discrepancies or disagreements in the analysis and interpretation of the information or in the transcripts themselves, with the interviewee having the final authority. The interviewee may withdraw his/her participation in the project no later than 2 weeks after the final "member checking process" by submitting a written request to the researcher and her doctoral committee (See Appendix -- for copy of consent form/research and data agreement). After that time all efforts will be made to negotiate any final discrepancies or disagreements between the researchers and the participants.

**Presentation of Data and Findings**

The final presentation of data and findings generated in this study will take several forms. The "stories" or indepth descriptions will be the most commonly used format and will be presented in a narrative format with direct quotes accentuating the language and perspective of the interviewees. A matrix for contrasting the sentient evaluators with each other will emphasize the major areas of interest in the study and the topics or issues that are raised in the data collection and analysis. A matrix for contrasting the role of the sentient with other more traditional roles outlined in the literature will emphasize communicating and reporting methods, involvement of various groups in the evaluation process, and ethical dilemmas faced by the evaluators, as well as other topics or issues emerging from the data. A matrix may be used to chart the first and second interview sessions to highlight the specific topics discussed and what additional information was gathered.

As a means of recording main themes, impressions, speculations, and implications evolving from the interviews, a modified "site analysis" form will be attached to each interview transcript, which may prove useful in tracking the credibility, dependability, and confirmability of the study. Finally a "member checking" form will be used to help obtain feedback from the interviewees on the credibility and accuracy of the "stories" or descriptions which are produced by the data analysis and writing sessions.
Appendix C

Informed Consent Form
INFORMED CONSENT FORM

RESEARCHER: Mary Elizabeth Piontek
Contact Address: 1728 Birchton, Portage MI 49002
616-327-7753/327-6154 h. 616-387-3791 w.

Affiliation: Doctoral Student - Measurement, Research, & Evaluation
Dept. of Educational Leadership
Western Michigan University
Kalamazoo MI 49008
616-387-
Doctoral Committee: Drs James Sanders and Zoe Barley
- Western Michigan University, John Seeley, FERA Inc.
Ann Arbor MI

STATEMENT OF PURPOSE

The overall purpose of this study is to look at the ways sentient evaluators involve stakeholders, clients, and audiences in the evaluation process and communicate their findings to these various groups and the ethical dilemmas that this role poses in (a) how they involve their stakeholders, clients, and audiences in the evaluation process (context) and (b) how they communicate, especially in terms of the language they use and the "voices" (perspectives) they attempt to capture (format/content).

DURATION OF PARTICIPATION

The research project is anticipated to take place from September 1993 through June 1994. The participants will be interviewed in November 1993, with follow-up interviews and data analysis feedback continuing through May 1994.

PROCEDURES

This study will utilize indepth semi-structured and unstructured ethnographic interviewing techniques on a select population of evaluators identified by 1) their own statements and/or published research, or 2) by their involvement in/presentations for the Qualitative Research, Women in Education, and Research on Evaluation Special Interest Groups of AERA, or 3) by their involvement in and presentations at American Evaluation Association Topical Interest Groups of Utilizing Evaluation and Qualitative Methods, or 4) by recommendations by practicing evaluators.

The interviews will be a combination of three types of interview instrumentation: informal conversational, interview guided, and standardized open-ended. The development, piloting, and final revision of the interview protocols involved using the information gathered through literature review, past evaluation work and research, and feedback from practicing evaluations concerning the accuracy, relevancy, and appropriateness of the language of and issues in the questions.

The following areas of inquiry are the major focus: model of evaluation used/method of inquiry or investigation; role of stakeholders/clients/audiences in evaluation phases (level of involvement in dev. of evaluation plan, data collection and analysis, writing of "reports", and communicating and reporting findings); communicating and reporting findings - perception of importance, individuals involved, methods used, and satisfaction with methods; ethics and social responsibility - personal views/values, extent to
which these enter into choice of evaluation projects, extent to which these enter into the evaluation process, extent to which they are articulated to audiences/clients/stakeholders, etc.

**RISKS/BENEFITS**

My intent is neither to promote or criticize this role, but rather to develop detailed descriptions about the role in terms of communicating and reporting efforts, the involvement of various groups in the evaluation process, and ethical considerations, in order to better understand its implications for the profession of evaluation.

Although all efforts will be made to assure the confidentiality of the information, it should be noted that it is not anticipated that the information collected during this research project will be "controversial or sensitive" or in any way will jeopardize the well-being, livelihood, or professional standing of the participants. By actively engaging the evaluators in dialogue, they can better understand themselves, their involvement with various groups in the evaluation process, ethical dilemmas, and the implications of their role for the profession of evaluation.

**CONFIDENTIALITY OF RECORDS/INFORMATION**

The researcher will use a portable tape-recorder with microphone to tape the interviews within a closed setting (private room or office). Only the researcher and the interviewee will be present during the interview session which will be approximately 45-90 minutes in length. All tapes produced during the interviews will be the property of the researcher and will be only in her possession before/during/after the interview sessions and research process. The tapes will be identified by the date of the interview, the setting, and the interview identification number (ID#). The identification number will be known only to the researcher and the interviewee so that transcripts and interpretive descriptions can be given to the individual interviewees for feedback and revision. The code will be a three-digit number (e.g., 003) assigned to the interviewee by the researchers and secured in the researcher's personal files so that only she can match the ID# to the interviewee's name.

The researcher will do all transcription of the interview tapes within a closed setting (her personal home office) so that no other individuals have access to the "raw"information. Any names or identifying information discussed on the tapes will be substituted for using the interviewee ID# or noted using dash marks or pseudonyms (e.g., Western Michigan University = ----University). The transcripts will be housed on computer discs accessible only to the researcher and identified only by the interviewee identification number. The transcripts will be stored in this format to facilitate the data analysis using ETHNOGRAPH (Qualis Research Associates, 1988), a qualitative data analysis software package.

The researcher will review the transcripts to develop a preliminary coding scheme. Printed copies of the transcripts will be sent to the individual interviewee. These printed copies will only be accessible to the researcher (and obviously the interviewee). The purpose of providing the interviewees with copies of their transcripts and drafts of analysis description or stories is to have them review, revise, confirm, or reinterpret the information.

The researcher and the interviewee will discuss any major discrepancies or disagreements in the analysis and interpretation of the
information or in the transcripts themselves, with the interviewee having the final authority.

PARTICIPATION

All participation is voluntary, and participants may withdraw their consent at any time during the research process continuing through June 1994 by contacting the researcher and/or her doctoral committee members.

<table>
<thead>
<tr>
<th>PARTICIPANT:</th>
<th>NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADDRESS</td>
</tr>
<tr>
<td></td>
<td>AFFILIATION</td>
</tr>
<tr>
<td></td>
<td>PHONE #s</td>
</tr>
</tbody>
</table>

SIGNATURE: ____________________________

DATE: ____________
Appendix D

Timetable With Notes
## TIME TABLE OF DISSERTATION ACTIVITIES

with notes/reflections on all activities

<table>
<thead>
<tr>
<th>Activity</th>
<th>Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final revision/writing of proposal (Ch.s 1-3, with consultation/feedback with committee chair)</td>
<td>June 28-August 31</td>
</tr>
<tr>
<td>Discussed general purpose of the proposal, possible data collection techniques, possible subjects and using a random comparison group, and vocabulary-language in the proposal during EDLD 695 Dissertation Seminar (Spring 93 Semester) -- other discussions with practicing evaluators, fellow graduate students, and work colleagues centered around the purpose of the proposal, possible subjects, and the vocabulary-language--ongoing conversations with MB, EG, RT, KW, PJ have kept me on track with literature review, methodology, and presentation of findings. Working on book proposal with RT, in some ways it is a very distracting project to be involved in, but it is helpful in the sense that it focuses on Com/Rep -- a subpart of my dissertation.</td>
<td></td>
</tr>
<tr>
<td>Proposal given to committee chair for discussion/feedback</td>
<td>Sept. 1-12</td>
</tr>
<tr>
<td>Discussed with Dr. Sanders on September, 1993 the general proposal, its purpose, the use of a random comparison group, and vocabulary-language.</td>
<td></td>
</tr>
<tr>
<td>Changes were made to the &quot;adjective&quot; used to describe the targeted evaluators and the use of the comparison group was added.</td>
<td></td>
</tr>
<tr>
<td>Discussion went very well--hope that entire process goes as smoothly!</td>
<td></td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Revision of proposal/committee meeting</th>
<th>Sept. 13-30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met with Committee (Dr.s Sanders, Barley, and Seeley) on September 28, 1993. Discussed Doctoral Program, Specialty-Comprehensive Exams, and proposal. Committee recommended minor changes to the discussion of the targeted evaluators, including adding detailed information on the &quot;definition/criteria&quot; of what makes up a &quot;sentient&quot; evaluator. Other suggestions included adding information to the literature review and changes in vocabulary-language to distinguish between various types and levels of roles. Noted that dissertation will need to discuss that some interviews took place face to face and some via the telephone. Ideas for thought included looking at parallels to action research, the current society-world situation (political, social, governmental) --does this evaluation worldview/framework reflect the times?, the Unit in the presentation of the data findings --sort of an overall composite, versus a focus on individuals. Discussed committee meeting with EG, MB, and RT. MB has been really helpful in thinking through my vocabulary, purpose, and focus.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Contact with possible interviewees, consent forms, set up dates/times for interviews</th>
<th>Oct. 4-27</th>
</tr>
</thead>
<tbody>
<tr>
<td>One target evaluator was in Kalamazoo, MI during October 1993 and consented to participate in the study. The interview was done in person at WMU during October 1993. The individual was given an overview of the study and signed a consent form. Both target (20) and random (20) groups were mailed a cover letter and an overview of the study during the first 2 weeks of October. Each person was contacted by phone and asked to participate. Ten Targeted and Five Random agreed. An additional 10 random persons were contacted by mail/phone, two were able to participate. Thirteen interviews were scheduled at AEA. One person (a target) was solicited for participation following a paper presentation and agreed to be interviewed = 14 total interviews. Four persons were unable to attend AEA and agreed to phone interviews. These were scheduled for convenience with their work schedules and to facilitate audio-taping of the interviews. Two were members of the target group, two of the random.</td>
<td></td>
</tr>
<tr>
<td>Get any final consent, feedback, and approval if necessary</td>
<td>Oct. 18-31</td>
</tr>
<tr>
<td>----------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>All forms for HSIRB approval were submitted in October, 1993. Complete approval was received prior to the interview process.</td>
<td></td>
</tr>
<tr>
<td>No additional approval was needed from my committee -- they approved the proposal on September 28, 1993.</td>
<td></td>
</tr>
<tr>
<td>Some participants submitted their consent form via mail prior to AEA. Others presented their forms before the interview took place. Spoke with MB and EG about HSIRB process -- both wanted information on my approach/level of detail.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interviews/data collection</th>
<th>AEA Nov. 3-6 in person December 1993-January 1994 for phone interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fourteen interviews were conducted at AEA over a four-day period. Eleven took place in my hotel room, one over breakfast in the hotel restaurant, and two in the conference lobby. All were audiotaped.</td>
<td></td>
</tr>
<tr>
<td>The interviews went very well. Only one minor problem with the remote microphone caused a small portion of one interview to be lost (this information was recreated from the interviewer's notes and by the interviewee adding information to the transcript). Each interview lasted between 30-60 minutes. All participants were enthusiastic and interested in the interview process and areas of focus. Some were more difficult to interview than others due to their natural conversation style or degree of introversion/extroversion, but each responded to all of my questions.</td>
<td></td>
</tr>
<tr>
<td>During the interviews, I took sketchy notes so that I could immediately reflect on the interview process and each participant's insights following each interview session. I rewound each tape and listened briefly to various sections to make certain that the tapes/recorded had not malfunctioned. One tape had a minor problem that was easily recreated by the interviewer and the interviewee once the transcript was produced.</td>
<td></td>
</tr>
<tr>
<td>Even at this stage I began to think about possible coding and ways to present the findings of the study: stories/composites, comparison-continuums, etc. Discussed with KW, RT, and MB.</td>
<td></td>
</tr>
</tbody>
</table>
**Transcription of Interviews**

<table>
<thead>
<tr>
<th>Date: Nov. 8-28, 1993</th>
<th>December 1993-January 1994</th>
</tr>
</thead>
</table>

The one interview that took place at WMU in October and the 14 interviews that took place at AEA were transcribed during the weeks of November 8 through November 28. At most two tapes could be transcribed each day -- most of the tapes took between 2.5 and 4.0 hours for this unexperienced transcriber! The transcripts were mailed on Monday November 29, 1993.

The phone interviews were transcribed within two days of each interview, to facilitate quick turn-around on feedback from the participants.

Thirteen participants returned their transcripts with corrections/additions between December 1993 and January 1994. These corrections/additions were made and these new transcripts were used for coding and analysis.

**Preliminary reading of transcripts for themes/categories/coding**

| Date: January 1994 |

As participants submitted their revised transcripts and using my interview notes, I developed the first-round of codes and thought about possible ways of grouping, discussing, and presenting the findings.

These were only preliminary ideas that were subsequently revised as the codes were applied/processed on the transcripts and findings emerged.

Discussed with PJ, RT, and MB. These people are really becoming my informal committee members. Usually we only casually discuss minor point or two, usually in connection with some other research they are doing, but it helps me stay on top of my dissertation, without taking up too much time or doing needless writing/revision at this point.

**Develop preliminary coding scheme**

| Date: January-February 1994 |

The first-round codes consisted of 21 codes developed from the interview protocol and the transcripts. These were used twice to code the transcripts to look for redundancy in the codes themselves and for pieces of the transcripts that lacked codes or necessitated the development of additional codes.

---

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
<table>
<thead>
<tr>
<th>Use coding scheme/Review-revision of coding scheme/Use of coding scheme</th>
<th>February 1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>The first-round codes were used twice, within a three-day period, to code the transcripts. Two codes were eliminated due to redundancy and the remaining 19 codes were grouped according to areas of focus/topic. Using the second-round codes all transcripts were recoded.</td>
<td></td>
</tr>
<tr>
<td>These 19 codes were entered into ETHNOGRAPH. This computer program enables the research to code individual lines or a series of lines in a document and then search each document (interview transcripts) using these codes.</td>
<td></td>
</tr>
<tr>
<td>I ran seven searches on the transcripts, utilizing ETHNOGRAPH's ability to search codes in the order determined by the researcher, based on the seven &quot;grouping&quot; of the codes I developed.</td>
<td></td>
</tr>
<tr>
<td>These &quot;groupings&quot; facilitated the reading of the data and the development of the findings, since similar topics/issues were near each other on the computer printouts, eliminating extra paper shuffling/sorting or the need to physically cut and paste the printouts.</td>
<td></td>
</tr>
<tr>
<td>Within each grouping I separated those responses/data that belonged to the target evaluators and those that belonged to the random group. This process went very smoothly, probably because I have used ethnograph before, so I could do most of the computer commands without too much of a struggle. I also read through the manual to get a sense of the most expedient way of organizing the searches and the print outs.</td>
<td></td>
</tr>
</tbody>
</table>
Analysis of coded data

Following the seven searches run on the transcripts, each "grouping" was put in a separate file folder. Each folder was read for consistency of the narrative information with the designated code (i.e., Were any lines or series of lines miscoded? Was any information confusing because it was taken out of context? or because pieces of the line was lost?, etc.). No major problems were found within any of the coded outputs.

One feature of ETHNOGRAPH that is particularly helpful during analysis is that for each search the computer software produces a list of those documents (interview transcripts) in which each code was not found. This facilitates going back to individual transcripts to see if there is indeed no data that matches that code or if such information was just overlooked by the researcher. On two codes I found areas where information had been overlooked or listed under another code. These portions of the transcripts were recoded and added to the respective file folder.

Once all of the coded outputs were read for consistency and accuracy, I began to list basic issues/statement from each respondent on a separate sheet of paper. From this basic list, I looked at the similarities/differences between the individual evaluators and the two groups (target and random) and began to brainstorm ways to present this information: Issue-lists/Areas of Focus Descriptions, on Continuums, as Contrasts-Counterpoints, as Composite Stories or overviews?

As always, using individual folders and highlighting quotes and key issues as I read through the printouts has saved a lot of time when I need to find a quote or identify which issues to emphasize in the write up.

Development of categories/themes

I utilized all of the preliminary notes and highlighting to expedite this process. As I worked through each folder (a minimum of three times) I refined and refocused how I might present each major area of focus. Although their are numerous possibilities -- I think that a combination of summary statements with direct quotes, visual depictions on continuums of where individuals sit, and an overall 'story' on the sentient --will produce a nice balance of specific to relationships to general overview in terms of perspective.

Went to College of Education Research presentation -- Pat Jenlink on Focus Groups --initially went to take notes for MF for her coursework/job, but found that the audience members discussed how to write up qualitative data, interview data, information from focus groups--it was helpful to her how others have thought through this process.

February - April 1994
<table>
<thead>
<tr>
<th>Revision/Further Refinement of categories/themes</th>
<th>March-April 1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>Again, as I work through the categories, I am slightly reformatting and rethinking the possibilities. It will be helpful when I get this entire section completed and can go back and review it 'enmass' and then brainstorm the implications--what does this data really tell me?</td>
<td></td>
</tr>
</tbody>
</table>

I had a brief meeting with Dr. Sanders concerning my Specialty Comps where we also discussed my dissertation--he brainstormed some ideas on research that could follow my dissertation (continuums or ratings for evaluators on their approaches, etc.). It helped me to talk about the fact that there really are no "black and white" distinctions between the target and random group -- it really is in a sense a series of continuums (yes, some of them quite artificial or strained) on which an individual may sit. Sometimes it is more the degree of intensity (for instance the degree the evaluator believes in social responsibility and who that touches) than the complete presence or absence of a quality or belief.

Have spoken with MB and KW about this--need to continue to refine this discussion.

<table>
<thead>
<tr>
<th>Development of &quot;stories&quot;/writing</th>
<th>March - April 1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is the most difficult piece, since my talent does not lie in creating semi-artistic, poetic, short-stories.</td>
<td></td>
</tr>
</tbody>
</table>

I am trying to combine the essence of each target, the complete discussion that we had (the transcript), the major issues and quotes that emerged, and the portrait that I created in the literature review--and how it is different/similar with other roles/random group.

This will have to be written at the end of the process--I need to think through the parts, before I can create the whole. Have discussed with MB and EG. They are encouraging--MB may create similar writing output in her research.

<table>
<thead>
<tr>
<th>Distribution of Draft Dissertation to Committee Chair and Revision of draft Subsequent distribution of Revised-Draft to dissertation committee</th>
<th>April 1994 May 1994</th>
</tr>
</thead>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Distribution of "stories" to interviewees for member checks

Draft findings were distributed to each participant so that feedback concerning the accuracy, consistency, and value of the findings could be solicited.

Each participant received two pieces or subsections of the findings to critique. Since the numerical codes given to each participant were altered in the presentation of the findings, I did not try to match an individual participant with his/her direct quotes that were presented in the findings (in otherwords I did not try to match interviewee originally coded as 00# to his/her comments that were designated as R0# or T0# depending on which group the interviewee is from).

No personal or controversial information was presented in the direct quotes so it is impossible for the participants to identify one another -- in fact since I re-numbered them in the presentation of the findings an individual may not recognize his/her own quotes.

Looking forward to their feedback--many of them use qualitative data, interesting to see what kind of advice/criticism they will have not only on what I presented as findings, but also how I presented the findings.

<table>
<thead>
<tr>
<th>Final stories</th>
<th>April-June 1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revision/feedback from committee</td>
<td>April - June 1994</td>
</tr>
</tbody>
</table>
Appendix E
Site Form
CONTACT SUMMARY FORM

Type of Contact: Who: Place: Date:

1. What were the main Themes, Impressions, Issues, Summary Statements about this contact? What was salient, interesting, illuminating, or important about this contact?

2. Explanations, Speculations about "what is going on"

3. What are possible alternative explanations?

4. What information did you get and fail to get on each of the target areas for this contact (main issues or topics)?

5. What are the next steps for data collection: followup questions, specific actions, general directions? What new or remaining target questions do you have in considering the next contact with this person?

6. What are the implications for revision, updating of coding scheme?
Appendix F

Member Checking Form
MEMBER CHECKING FORM

<table>
<thead>
<tr>
<th>Themes/Stories/Issues from Data Analysis</th>
<th>Accuracy, Trustworthiness</th>
<th>Feedback, explanation, revision, etc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please list the first few words of the quotation of the title/subtitle of the section you are referring to...</td>
<td>Please give your reactions...</td>
<td>Please provide any comments, suggestions, criticisms, and feedback...</td>
</tr>
</tbody>
</table>

Please return to Mary E. Piontek, 1728 Birchton, Portage MI 49002. (616-327-7753) via enclosed SASE. Thank you. If you need more writing space, please use the back of this page.
Appendix G

Matrix of Evaluation Roles
<table>
<thead>
<tr>
<th>ROLE CHARACTERISTICS</th>
<th>EVALUATOR ROLE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EXP</td>
</tr>
<tr>
<td>Positivism</td>
<td></td>
</tr>
<tr>
<td>Objectivity</td>
<td></td>
</tr>
<tr>
<td>Transfer of power</td>
<td></td>
</tr>
<tr>
<td>Evaluation</td>
<td></td>
</tr>
<tr>
<td>Utilization</td>
<td></td>
</tr>
<tr>
<td>Grounded in</td>
<td></td>
</tr>
<tr>
<td>participant</td>
<td></td>
</tr>
<tr>
<td>experiences</td>
<td></td>
</tr>
<tr>
<td>Holistic</td>
<td></td>
</tr>
<tr>
<td>Historical</td>
<td></td>
</tr>
<tr>
<td>Comparative</td>
<td></td>
</tr>
<tr>
<td>Multiple Perspectives</td>
<td>X S</td>
</tr>
<tr>
<td>Observer</td>
<td></td>
</tr>
<tr>
<td>Collaborator</td>
<td></td>
</tr>
<tr>
<td>Translator</td>
<td></td>
</tr>
<tr>
<td>Promotes</td>
<td></td>
</tr>
<tr>
<td>participant</td>
<td></td>
</tr>
<tr>
<td>ownership</td>
<td></td>
</tr>
<tr>
<td>Teacher</td>
<td></td>
</tr>
<tr>
<td>Program efficiency</td>
<td></td>
</tr>
<tr>
<td>Journalist</td>
<td></td>
</tr>
<tr>
<td>Emancipatory</td>
<td></td>
</tr>
<tr>
<td>Self-reflection</td>
<td></td>
</tr>
<tr>
<td>Theoretical</td>
<td></td>
</tr>
<tr>
<td>Artistic</td>
<td></td>
</tr>
<tr>
<td>Challenges</td>
<td></td>
</tr>
<tr>
<td>beliefs</td>
<td></td>
</tr>
</tbody>
</table>

Note: The abbreviation for the evaluation roles are as follows EXP = Expert, PM = Power Merchant, INT = Interpreter, MED = Mediator/Broker, FAC = Facilitator, CON = Consultant/Facilitator, PC = Personal Coach, INV = Investigator, TE = Theorist/Emancipator, CC = Connoisseur/Critic, SC = Social Critic/Critical Ethnographer. Where an S appears in the matrix boxes, this designates traits that these roles share with the definition of sentient evaluators presented in this study.
Appendix H

Coding Book
OVERVIEW

The following codes were developed for use on the data collected in connection with my research on "sentient" evaluators. This set of codes was developed to facilitate the analysis of 19 transcribed interviews (completed between 10/93-1/93). For each code a brief definition, descriptor, or example is given. Where appropriate, a specific interview question or series of questions is linked to a particular code.

It should be noted however that the interviewee may have discussed any of these topics during the course of the interview and not necessarily only in concert with a directed interview question and thus the coder must be alert to using all codes at any given point in the transcribed data.

Any given sentence, paragraph, section of paragraphs, etc. can contain numerous codes (in other words codes can overlap). The coder needs to keep this in mind and not be concerned in oversimplifying the coding of data by using only one code for a section or subsection of the data.

CODES

SETTNG = Settings/areas where person has conducted or is currently conducting evaluation: (Let me first ask you a few questions about your background in evaluation: - In what settings/professions/areas have you conducted evaluations? - In what area(s) do you most frequently work? - In what area(s) do you plan to work in the future?)

PRFROL = Professional role(s) of the person, may not see him/herself primarily as an evaluator (- In what other professions do you work? - i.e., higher ed., government, etc. What do you consider as your main professional role?)

YEARS = Years experience conducting evaluations (- For how many years have you been working in the field of evaluation?)

PROCES = Typical evaluation process/inquiry method that the person uses. Some persons may respond that they do not have one, if so such a response should also be coded as PROCES (- Could you describe a typical evaluation process that you frequently use? - Do you have a particular model of evaluation or method of inquiry that you most frequently use? - To what extent do your views/values enter into the evaluation process?)

CLCOM = Client communication is how the person discusses
the evaluation process/model that they use as well as all other forms of communication (formal and informal, verbal and written) used with the clients/audiences/stakeholders. Do you communicate this to your clients/audiences/stakeholders? When? How? - How much of your total evaluation time do you spend on com/rep efforts? - Who are the individuals involved? - To what extent do your views/values enter into the evaluation process? - Do you articulate them to your clients/audiences/stakeholders? How?)

CLTINV = Client/s/a involvement describes how the person involves these groups in the evaluation process (from planning the evaluation to writing formal reports at the conclusion of the evaluation) - What role do clients play in your evaluation process? - audiences? - stakeholders? - Are clients/audiences/stakeholders involved in the development of the evaluation plan? How? - in the collection of data? ...the analysis of data? - in the presentation of findings (writing reports/presentations) - in dissemination?)

PURPOS = Purpose of evaluation involves all purposes, goals, intents that the person ascribes to evaluation processes (- What do you see as the main purpose of evaluation?)

BCKVAL = Background/values describes all training, education, evaluation, and professional experiences discussed by the person. The focus of the interview was on the effect of the person's background on how they conduct evaluation and what they see as the purpose of evaluation, but all discussion of the background/values should be coded as BCKVAL, even if it overlaps with other codes (e.g. SETTNG, PRFROL) - How does this view reflect your values and background? - To what extent do your views/values enter into the evaluation process? - Do you articulate them to your clients/audiences/stakeholders? How?)

SOCRES = This code is used where the person discusses the Social Responsibility/Role in Society of evaluation. This code may overlap with PURPOS, FTISSU, and BCKVAL at times. (- What do you see as the role of the profession of evaluation in terms of society/social responsibility?)

COLLEA = This code should be used when the person discusses his or her colleagues, both in the field of evaluation and outside of the field. Each person was asked about his/her colleagues in connection with their views of the purpose of evaluation and its social responsibility, but all mentions of colleagues or peers should be coded as COLLEA (- Do you think that this differs from most of your peers/colleagues?)
CRIMPT = This code refers to communicating and reporting efforts used by the evaluator and the amount of importance he or she places on these efforts. This code may overlap with CRACTV, CRSATF, and CRFDBK. (- How much importance do you place on communicating and reporting efforts?)

CRSATF = This code refers to the person's satisfaction with the communication efforts that he or she uses in an evaluation context. (- How satisfied are you with these activities?)

CRACTV = This code is used to designate all communicating and reporting activities, whether formal or informal, verbal or written, used in an evaluation context. (- How much of your total evaluation time do you spend on com/rep efforts? - Who are the individuals involved? - What kind of com/rep activities do you use?)

CRFDBK = All feedback the person has received from clients/audiences/stakeholders involving any communication efforts. (- What feedback have you gotten from clients, audiences, stakeholders concerning these activities?)

CHOICE = The code refers to the choice the persons make in terms of evaluation clients/contracts/experiences. (Earlier you said ...was the main purpose of evaluation. Does this view/value, enter into your choice of evaluation projects? Have you ever turned down a client because their evaluation purposes conflicted with your views? Did you try to resolve these conflicts? How?)

ETHDIL = All ethical dilemmas/conflicts encountered by the person in an evaluation context should be coded as ETHDIL as well as any discussions of conflict of personalities, philosophies, methods, ideas or values. (- Do you think that the way that you conduct evaluations or your view on the purpose of evaluation poses any particular ethical dilemmas? Could you give me an example?)

ETHRES = This code reflects any discussion of resolutions the person has used in an evaluation context for ethical dilemmas/conflicts. (-Do you have any particular methods that you use to help resolve or soften these dilemmas?)

FTISSU = All issues in terms of evaluation that the person sees as needed to be discussed or focused on by the professional evaluation community or associations. (- What issues do see as most salient for the future of evaluation?)

FTACTV = Future activities to be used to address these evaluation issues should be coded as FTACTV. This might include training of future evaluators, association...
activities, professional discussions, etc. (- Do you think that your views/methods/activities in evaluation can help focus/expand these issues?)

EXAMPLES OF USING THESE CODES

COLLEA: "Yes, I do think that it does. I mean that I think that this is a much broader view, what I suggest is a much larger view. I think that most of them would tend to see themselves as doing a program, approaching a client, a sponsor of an evaluation who would be the client..."

SETTING: "A whole host of them. You know I am with extension. I do program evaluation work with cooperative extension, so I do it in the non-formal educational setting. In all the different program areas as we call them in extension. We do programming in agriculture and natural resources. We do programming in Home Ec., some call it family living. I've done evaluations of 4-H youth..."

CLTINV: "Well they don't recognize usually that they are being involved. I am not one who tries to hammer out a conceptual plan that formally takes into consideration the avowals or the fears of stakeholders, but I am trying to be sensitive to them, and as I know them let them influence what my team is paying attention to. This includes a deliberate effort to seek out and respond to people who have contrary ideas to the expected plan..."
Appendix I

HSIRB Protocol Approval
Date: October 8, 1993
To: Mary Elizabeth Piontek
From: M. Michele Burnette, Chair
Re: HSIRB Project Number 93-10-03

This letter will serve as confirmation that your research project entitled "Responsive evaluators: Their role in communicating with and involving stakeholders in the evaluation process" has been approved under the exempt category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

You must seek reapproval for any changes in this design. You must also seek reapproval if the project extends beyond the termination date.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: October 8, 1994

c: Sanders, Ed. Leadership
WESTERN MICHIGAN UNIVERSITY
HUMAN SUBJECTS INSTITUTIONAL REVIEW BOARD (HSIRB)
HUMAN SUBJECTS APPROVAL FORM

RESEARCH MAY NOT BEGIN UNTIL THE PROTOCOL HAS BEEN REVIEWED AND APPROVED BY THE HUMAN SUBJECTS INSTITUTIONAL REVIEW BOARD, WHICH MEETS ON A REGULAR MONTHLY BASIS. PROTOCOLS MUST BE RECEIVED BY RESEARCH AND SPONSORED PROGRAMS AT LEAST SEVEN DAYS PRIOR TO A REGULARLY SCHEDULED MEETING IN ORDER TO BE ACTED ON AT THAT MEETING. THE FORM MUST BE TYPEWRITTEN, EXCEPT FOR SIGNATURES.

PRINCIPAL INVESTIGATOR* Mary Elizabeth Piontek
DEPARTMENT Educational Leadership
Office address: 3312 Sangren Hall Office Phone: 616-387-3879
Home Address: 1728 Birchton Portage Michigan 49002
Home Phone: 616-327-7753 (Zip Code)

PROJECT TITLE: Responsive Evaluators: Their role in communicating with and involving stakeholders in the evaluation process.

PROPOSED PROJECT DATES From 11/93 To 6/94
SOURCE OR POTENTIAL SOURCE OF FUNDING none
APPLICATION IS New X Renewal

Protocols for projects extending beyond one year from date of HSIRB approval must be submitted annually for renewal.

If this proposal is approved by the Institutional Review Board, the Principal Investigator agrees to notify the HSIRB in advance of any changes in procedures which might be necessitated. If, during the course of the research, unanticipated subject risks are discovered, this will be reported to the IRB immediately.

Mary E. Piontek 9/30/93
P.I. Signature Date

*If the Principal Investigator is a student, complete the following:
Undergraduate Level Research Graduate Level Research Doctoral - Ph.D.
Faculty Advisor Dr. James Sanders Telephone 616-387-5895
Department Educational Leadership
Advisor Signature Date

Rev. 8/92 All previous forms are obsolete and should not be used.
ABSTRACT: Briefly describe the purpose, research design, and site of the proposed research activity. The purpose of this study is to look at the ways evaluators involve stakeholders in the evaluation process and communicate their findings. An evaluative evaluator takes an active role in promoting the program that she is investigating and deliberately seeks to represent underrepresented stakeholders. This study will use structured/semi-structured/unstructured interview questions for use in individual interview sessions with no more than 30 professional evaluators at professional conferences and follow-up phone interviews. The interview transcripts will be coded, organized into themes/categories, and written into descriptive narrative stories to fully describe the role of the evaluative evaluator in terms of stakeholder involvement, communicating and reporting efforts, and the unique dilemmas of the role.

BENEFITS OF RESEARCH: Briefly describe the expected or known benefits of the research. While factors about reporting/communicating evaluation findings have been addressed in the literature, there has been no systematic, comprehensive treatment of the topic in terms of the issues currently faced by evaluators in a variety of settings/roles. Little has been written about the role of the evaluative evaluator and the ethical situations encountered by these evaluators. This study will contribute to this body of knowledge, as well as address some of the more salient issues of discussion in evaluation today: ethical standards, collaboration, alternative comm/rep efforts. By actively engaging the evaluators in dialogue, they can better understand themselves.

CHARACTERISTICS OF SUBJECTS: Briefly describe the subject population (e.g., age, sex, prisoners, people in mental institutions, etc.). Also indicate the source of subjects. Professional Evaluators, Age 21+, male/female, United States residents, members of the American Evaluation Association and/or the American Educational Research Association. Selected by their 1) PROFESSIONAL RESEARCH OR PUBLICATIONS, 2) PRESENTATIONS AT AEA and/or AERA and various interest groups of these associations, and/or 3) RECOMMENDATIONS BY PRACTICING EVALUATORS.

SUBJECT SELECTION: How will the subjects be selected? Approximately how many subjects will be involved in the research? (Attach advertisement for subjects [Cover letters used in survey research are equivalent to advertisements. Scripts are equivalent to oral solicitation procedures].) Not more than 30 professional evaluators will be included in the study. The potential subjects will be contacted by mail (see Apx A) to solicit their involvement in the study. A follow-up phone call will be made one week later to solicit verbal agreement to participate. Those who give verbal agreement will be sent a consent form (Apx B) along with an abstract of the study, to sign and return. The subjects may withdraw their consent at any time during the data collection/analysis/and final written descriptive process.
VULNERABLE SUBJECT INVOLVEMENT (Fill out if applicable)

Research involves subjects who are (check as many as apply)

1. ___ Children (any subject under the age of 18) Approximate age ___
2. ___ Mentally retarded persons
3. ___ Mental health patients
4. ___ Check if institutionalized
5. ___ Prisoners
6. ___ Pregnant women
7. ___ Other subjects whose life circumstances may interfere with their ability to make free choices in consenting to take part in research:

(Describe)

LEVEL OF REVIEW

To determine the appropriate level of review, refer to WMU Policy Guidelines for categories of exempted research (Appendix B).

X Exempt (Forward the original application to the Chair of the Department for a cover letter, then forward to HSIRB Chair via RSP)

Subject to Review (Forward original application plus 9 copies to HSIRB Chair via RSP)

BLOOD PRODUCTS INVOLVED □ No □ Yes, attachment appended

If your research involves the collection of blood or blood products, then pick up and complete an addendum (HSIRB Collection of Blood and Blood Products Form).

PLEASE TYPE THE REQUESTED PROTOCOL INFORMATION ON THE FOLLOWING PAGES OR USE THE ELECTRONIC FILE AVAILABLE. You may attach additional sheets as necessary and reference the appropriate page.
RISKS TO SUBJECTS: Briefly describe the nature and likelihood of possible risks, or discomfort (e.g., physical, psychological, social) as a result of participation in the research. The subjects will be given copies of their interview transcripts, draft analysis codes and themes/descriptions, preliminary written descriptions, and final written descriptions to revise and edit. Neither their names, professional affiliation, or other professional will be used. There should be no risk to the subjects in terms of their well-being, livelihood, or professional standing.

PROTECTION FOR SUBJECTS: Briefly describe measures taken to protect subjects from possible risks, or discomfort if any. Please see attached sheet for complete discussion.

CONFIDENTIALITY OF DATA: Briefly describe the precautions that will be taken to ensure the privacy of subjects and confidentiality of information. Be explicit if data are sensitive. Describe coding procedures for subject identification numbers. Please see attached sheet for complete discussion.

INSTRUMENTATION: Questionnaires, interview schedules, data collection instruments, should be identified. Attach a copy of what will be used in this project. Coding sheets for video-tape or audio-tape data collection procedures are required. Interview protocols, a computer data analysis program (ETHNOGRAPH), a portable tape-recorder, a tentative research-interview schedule—all instrumentation is discussed on the attached sheet. See Appendix C for instrumentation documents.

INFORMED CONSENT: For further information on writing consents (assents not covered), see the book Informed Consent by T. M. Grundner, on reserve at Waldo Library. Attach a copy of the informed consent and assent (if applicable). Each subject should also be given a copy. Please see attached sheet for complete discussion.
Mary E. Piontek  
1728 Birchton Portage MI 49002  616-327-7753

Human Subjects Institutional Review Board Application continued:

PROTECTION OF SUBJECTS

All tapes produced during the interviews will be the property of the researcher and will be only in her possession before/during/after the interview sessions and research process. The tapes will be identified by the date of the interview, the setting, and the interview identification number (ID#).

The identification number will be known only to the researcher and the interviewee so that transcripts and interpretive descriptions can be given to the individual interviewees for feedback and revision. The code will be a three-digit number (e.g.; 003) assigned to the interviewee by the researchers and secured in the researcher's personal files so that only she can match the ID# to the interviewee's name.

Although all efforts will be made to assure the confidentiality of the information, it should be noted that it is not anticipated that the information collected during this research project will be "controversial or sensitive" or in any way will jeopardize the well-being, livelihood, or professional standing of the participants.

The interviewee may withdraw his/her participation in the project by submitting a written request to the researcher and her doctoral committee (See Appendix B for copy of consent form/research and data agreement). All efforts will be made to negotiate any final discrepancies or disagreements between the researchers and the participants.

CONFIDENTIALITY OF DATA

The researcher will use a portable tape-recorder with microphone to tape the interviews within a closed setting (private room or office). Only the researcher and the interviewee will be present during the interview session which will be approximately 45-90 minutes in length.

All tapes produced during the interviews will be the property of the researcher and will be only in her possession before/during/after the interview sessions and research process. The tapes will be identified by the date of the interview, the setting, and the interview identification number (ID#).

The identification number will be known only to the researcher and the interviewee so that transcripts and interpretive descriptions can be given to the individual interviewees for feedback and revision. The code will be a three-digit number (e.g.; 003) assigned to the interviewee by the researchers and secured in the researcher's personal files so that only she can match the ID# to the interviewee's name.
Although all efforts will be made to assure the confidentiality of the information, it should be noted that it is not anticipated that the information collected during this research project will be "controversial or sensitive" or in any way will jeopardize the well-being, livelihood, or professional standing of the participants.

**INSTRUMENTATION**

The researcher will use a portable tape-recorder with microphone to tape the interviews within a closed setting (private room or office). Only the researcher and the interviewee will be present during the interview session which will be approximately 45-90 minutes in length.

The interview transcripts, and any other documents/written materials, will be analyzed using an emergent design, in which a series of reflective and evolving data analysis phases will take place in order to capture the varying themes, categories, and/or "stories" that emerge from the data.

The analysis of the interview transcripts and other documents/written materials will follow the same basic design presented below in Figure 1.

**FIGURE 1**

**DATA ANALYSIS PROCESS**

Interview → Transcription → feedback from interviewees on transcripts → Preliminary reading and identification of themes → Development of preliminary coding scheme → Use of coding scheme on transcripts or documents → Analysis of coded data to identify themes or categories → Review and revision of coding scheme → Use of coding scheme on transcripts or documents → Analysis of coded data → Development of categories and themes → Development or writing of "stories" or descriptions → Distribution of "stories"/descriptions to interviewees for 'member checks' → may lead to

Second round of interview/followup → Use of coding → Revision of "stories" → 'member checks' → Final "stories" on communicating/reporting, collaboration/relationships/involvement, ethical considerations, and other categories that might emerge.

The researcher will do all transcription of the interview tapes within a closed setting (her personal home office) so that no other individuals have access to the "raw"information. Any names or identifying information discussed on the tapes will be substituted for using the interviewee ID# or noted using dash marks or pseudonyms (e.g.; Western Michigan University = —University).

The transcriptions will be housed on computer discs accessible only to the researcher and identified only by the interviewee identification number. The transcripts will be stored in
this format to facilitate the data analysis using ETHNOGRAPH (Qualis Research Associates, 198), a qualitative data analysis software package.

The researcher will review the transcripts to develop a preliminary coding scheme. Printed copies of the transcripts will be sent to the individual interviewee. These printed copies will only be accessible to the researcher (and obviously the interviewee). The purpose of providing the interviewees with copies of their transcripts and drafts of analysis description or stories is to have them review, revise, confirm, or reinterpret the information.

The researcher and the interviewee will discuss any major discrepancies or disagreements in the analysis and interpretation of the information or in the transcripts themselves, with the interviewee having the final authority.

The final presentation of data and findings generated in this study will take several forms. The "stories" or in-depth descriptions will be the most commonly used format and will be presented in a narrative format with direct quotes accentuating the language and perspective of the interviewees.

A matrix for contrasting the evaluators will emphasize the major areas of interest in the study and the topics or issues that are raised in the data collection and analysis (See Appendix D). A matrix for contrasting the role of the responsive with other more traditional roles outlined in the literature will emphasize communicating and reporting methods, involvement of various groups in the evaluation process, and ethical dilemmas faced by the evaluators, as well as other topics or issues emerging from the data (See Appendix D). A matrix may be used to chart the first and second interview sessions to highlight the specific topics discussed and what additional information was gathered (See Appendix D).

As a means of recording main themes, impressions, speculations, and implications evolving from the interviews, a modified "site analysis" form will be attached to each interview transcript, which may prove useful in tracking the credibility, dependability, and confirmability of the study (See Appendix D).

Finally a "member checking" form will be used to help obtain feedback from the interviewees on the credibility and accuracy of the "stories" or descriptions which are produced by the data analysis and writing sessions (See Appendix D).

INFORMED CONSENT

The interviewee may withdraw his/her participation in the project by submitting a written request to the researcher and her doctoral committee (See Appendix B for copy of consent form/research and data agreement). All efforts will be made to negotiate any final discrepancies or disagreements between the researchers and the participants.
BIBLIOGRAPHY


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


