



Western Michigan University
ScholarWorks at WMU

Dissertations

Graduate College

4-2000

A Study of the Relationship Between Epistemic Style and Evaluation Practice

Deanna Drazé
Western Michigan University

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



Part of the Educational Assessment, Evaluation, and Research Commons

Recommended Citation

Drazé, Deanna, "A Study of the Relationship Between Epistemic Style and Evaluation Practice" (2000).
Dissertations. 1447.

<https://scholarworks.wmich.edu/dissertations/1447>

This Dissertation-Open Access is brought to you for free and open access by the Graduate College at ScholarWorks at WMU. It has been accepted for inclusion in Dissertations by an authorized administrator of ScholarWorks at WMU. For more information, please contact wmu-scholarworks@wmich.edu.



**A STUDY OF THE RELATIONSHIP BETWEEN EPISTEMIC STYLE AND
EVALUATION PRACTICE**

by

Deanna Draz

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 Studies

Western Michigan University
Kalamazoo, Michigan
April 2000

A STUDY OF THE RELATIONSHIP BETWEEN EPISTEMIC STYLE AND EVALUATION PRACTICE

Deanna Draz, Ph.D.

Western Michigan University, 2000

The purpose of this study was to examine the differences between personal epistemic styles of professional evaluators, the influence those styles have on evaluation practice, and to determine whether a typology of evaluation practice can be extrapolated from a study of epistemic style. Epistemic style was defined as how one determines what is true (or what constitutes knowledge). The intent was to provide empirical evidence of the relationship between theory and practice in the hope that it will contribute to an increased awareness of the evaluator “as an instrument” through which data is filtered.

The Psycho-Epistemological Profile (Royce, Mos, and Learsley, 1975) was selected to measure dominant epistemic styles among professional evaluators. A reliability study of the instrument was conducted using a sample of evaluators from the Michigan Association of Evaluators. From this group, three evaluators were chosen for semi-structured interviews from each of the dominant epistemic styles represented in the sample. The represented epistemic styles were rational, metaphoric, and eclectic (those evaluators not having a dominant epistemic style).

Content analysis of the interviews revealed that the major differences between dominant epistemic styles were not so much in how evaluators determined purpose or use of evaluation but more in the processes involved in evaluation. These processes were (1) roles that evaluators, clients, and stakeholders play and (2) how analysis of

data was implemented. There were also differences in theoretical influences among epistemic styles.

From the interview data, a typology of evaluation practice was extrapolated. The typology was organized around four themes: (1) theory of action, (2) attitudes toward data, (3) focus on process and outcomes, and (4) theory and practice. The major difference was that the Metaphorics acted as change agents, the Rationals as educators, and the Eclectics as facilitators of improvement. All of these are related to a primary purpose of evaluation chosen by the three styles: to enable program staff to make changes that improve program effectiveness.

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.

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.

**Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600**

UMI[®]

UMI Number: 9963063

Copyright 2000 by
Draze, Deanna

All rights reserved.

UMI[®]

UMI Microform 9963063

Copyright 2000 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

**Copyright by
Deanna Drazé
2000**

ACKNOWLEDGMENTS

The writing of my dissertation was not a solo event. It involved many others in terms of support, encouragement, love, and direction. I thank my husband, Jim for his support, love, and advice. I appreciate all the hours he patiently listened to my ideas, giving me feedback when it was needed. I thank my friends Cynthia H., Cynthia P., Sharon, Kate, Lorraine, and Chialin for their enthusiastic encouragement. I thank my colleagues at SAMPI who supported my work. I thank my committee members Dr. Jianping Shen, Dr. Van Cooley, and Dr. Susan Edgerton for their assistance and encouragement. I am grateful to Dr. Uldis Smidchens, for his guidance and encouragement throughout my graduate years at Western Michigan University. Each of these persons played an important role in helping me accomplish my goal.

“Little acts of kindness which we render to each other in everyday life, are like flowers by the wayside to the traveler: they serve to gladden the heart and relieve the tedium of life’s journey.” Eunice Bathrick

Deanna Draz

TABLE OF CONTENTS

ACKNOWLEDGMENTS	ii
LIST OF TABLES	ix
LIST OF FIGURES.....	x
CHAPTER	
I. INTRODUCTION.....	1
Statement of the Problem.....	1
Theories of Knowledge and Evaluation Practice.....	1
Influences on Research and Evaluation Practice	1
Epistemic Style	2
Influence of Epistemic Style.....	2
Epistemic Style and Evaluation	3
Rationale for the Study	5
Epistemic Style as a Foundation for Evaluation Practice.....	5
Theory-Laden Observations	5
Consciousness of Personal Theories	6
Lack of Empirical Evidence	6
Demystifying of Theoretical Orientations.....	7
Personal Framework of the Study	8
Organization of the Study	8
II. LITERATURE REVIEW.....	10
Organization of the Review	10

Table of Contents—continued

CHAPTER

Influences on Evaluation Practice	10
Intervening Variables	10
Theoretical Orientations	10
Political Context.....	12
Culture of Organization and Community	13
Time and Budget	14
Gender and Race	15
Experience of Evaluator and Client	16
Relationship of Evaluator and Client	17
Definitions of Theory and Epistemic Style.....	18
Definition of Theory.....	18
Definition of Epistemic Style.....	19
Theories of Knowledge	20
Epistemic Styles and Behavior.....	21
Epistemic Styles and Evaluation Practice.....	22
Epistemic Styles and Evaluation Outcomes.....	23
Standardization as Coercion	23
Evaluation Conveying New Understandings.....	24
Typologies of Evaluation Practice.....	26
Categories of Practice.....	26
Accreditation and Accountability.....	26

Table of Contents—continued

CHAPTER		
	Emancipation	28
	Expertise	28
	Management.....	29
	Naturalistic and Participatory	30
	Theory-Driven.....	32
	Rationale for Study.....	33
	Evaluation Use.....	33
	Observations Are Theory-Laden.....	34
	Awareness of Evaluator as Instrument.....	35
	Multiple-Perspectives	36
	Summary	38
III. METHODOLOGY		39
	Differences in Epistemic Styles Among Professional Evaluators	39
	Selection of Instrument.....	39
	Description of Psycho-Epistemological Profile.....	39
	Population and Sample Selection for the PEP	47
	Administration of the Instrument.....	47
	Analysis of Instrument	48
	Reliability of the PEP With Professional Evaluators.....	49
	Relationship of Epistemic Style to Evaluation Practice	50
	Description of Methods	50
	Development of the Interview Protocol	51

Table of Contents—continued

CHAPTER		
	Selection of Sample for Interviews	52
	Interview Process	53
	Interview Data-Analysis Overview	54
	Establishment of Reliability and Validity of Interviews	56
	Summary	56
IV. DATA ANALYSIS		57
	Personal Style	57
	Relationship Between Epistemic Style and Evaluation Practice.....	60
	Research Question #1	60
	Differences Among Epistemic Styles.....	61
	Categories of No Difference	72
	Summary of Differences and Similarities in Evaluation Practice	77
	Epistemic Style as a Foundation for Evaluation Practice	78
	Research Question #2	78
	Epistemic Typology of Evaluators.....	79
	Definition of Themes Within the Typology	79
	The Typology.....	81
	Summary of Analysis.....	99
	Differences in Evaluation Practice.....	99
	A Typology of Evaluation Practice.....	101
V. SUMMARY AND DISCUSSION.....		103

Table of Contents—continued

CHAPTER

Major Findings	103
Difference in Practice	103
Epistemic Style as a Foundation for Evaluation Practice	103
Implications for Practice	104
Validity of Qualitative Research and Evaluation	104
Utility for Evaluators, Clients, and Stakeholders	104
Education of Future Evaluators	105
Organization of the Evaluation Team	106
Role and Nature of Evaluation	107
Limitations of the Study	108
Influence of Organizational Culture	108
Small Sample	109
Espoused Views	109
Lack of Empirical Style	109
Sole Author	110
Suggestions for Further Research	110
Expansion of Interview Protocol	110
Larger Sample	111
Ethnographic Methods	111
Educational Researchers	111
School Administrators and Teachers	112
Multiple Perspectives	112

Table of Contents—continued

CHAPTER

Conclusion.....	112
Link Between Epistemic Style and Evaluation Practice	112
Utility of Multiple Perspectives	113
Theory-Laden Observations	113
Practical Problems.....	114
The Future.....	114

APPENDICES

A. Autobiographical Bias	116
B. Human Subjects Institutional Review Board Approval.....	125
C. Informed Consent Form	127
D. Letters to Evaluators Requesting Participation	129
E. Permission to Use Psycho-Epistemological Profile.....	133
F. Interview Protocol.....	135
G. Case Study.....	139
H. Table 9: Interview Responses.....	141
BIBLIOGRAPHY	150

LIST OF TABLES

1. The Basic Paths of Knowledge	41
2. Correlations of Epistemic Style to Other Constructs	46
3. Reliability of Psycho-Epistemological Profile	49
4. Descriptions of Interview Sample	53
5. Primary Roles of Evaluators	65
6. Summary of Differences and Similarities of Evaluation Practice	78
7. Differences in Evaluation Practice by Epistemic Style	100
8. Typology of Evaluation Practice by Epistemic Style	101
9. Interview Responses	141

LIST OF FIGURES

1. Components of Meaning.....	55
-------------------------------	----

CHAPTER I

INTRODUCTION

Statement of the Problem

Theories of Knowledge and Evaluation Practice

What is truth? How is knowledge to be determined? These are critical questions for the professional evaluator. It seems logical that if evaluators hold different criteria for establishing knowledge or truth, the outcomes of their evaluations may differ even when evaluating the same program. However, does the evaluator's epistemic style (theory of knowledge) actually provide the foundation for evaluation practice? This is the central question for this study. The answer to this question addresses the core problem examined in this study: program evaluations generally are presented from the personal theoretical orientation of the evaluator but we do not have evidence of how theories of knowledge affect evaluation practice.

Influences on Research and Evaluation Practice

It has been well documented that there are many influences on the design of research methodologies such as theoretical orientations (Barger & Duncan, 1990; Marshall & Rossman, 1995; Patton, 1997), political context (House, 1993; Punch, 1994), culture or organization and community (Guba & Lincoln, 1989; Royce, 1964a), available time and budget (Worthen & Sanders, 1987), age and gender (Coan, 1979; Denzin & Lincoln, 1994; Royce and Mos, 1975), relationship of client and researcher (Bloom, 1998; Worthen & Sanders, 1987), and experience of the

researcher (Miles & Huberman, 1994; Worthen & Sanders, 1987). Denzin and Lincoln claim that it is an interactive process of these different factors among both the researched and the researcher (1994). These differences among researchers fuel the debate between the quantitative researchers from the positivist tradition and the qualitative researcher from the postpositivist tradition (Eisner & Peshkin, 1990). The differences are manifested also in the field of evaluation.

The field of evaluation has a close link to research methodology. A current definition of the term evaluate is “to determine the significance, worth, or condition of usually be careful appraisal and study” (Merriam-Webster’s Collegiate Dictionary, 1997, p. 401). This careful appraisal and study is accomplished by the use of research methodologies.

Epistemic Style

In the discussion of influences on research methodologies or evaluation practice, one’s personal epistemic style is rarely considered (Daly, 1997). Epistemic style is defined as how one determines what is true (or what constitutes knowledge) (Royce, 1964a). Epistemologies provide much of the justification underlying particular methodologies (Swandt, 1997b).

Rather than conceiving of the differences between so-called qualitative and quantitative inquiry in terms of tools and methods, students of qualitative inquiry might be better served by examining the differences between epistemologies, for example, the epistemologies of empiricism and hermeneutics. (Swandt, 1997b, p.39)

Influence of Epistemic Style

Epistemic style is personal and can be a process that is a function of beliefs and influences the selection of evidence, the construction of concepts, and forms a

pattern of relevancies, values, and priorities (Daly, 1997; Denzin, 1994; Royce, 1964a). Therefore, it seems that attention needs to be focused on understanding the thinking, reflecting processes of evaluators that rest on an evaluator's theory of truth and knowledge (epistemic style).

In his book, The Structure of Scientific Revolutions, Thomas Kuhn reversed the traditional empirical account that observations grant meaning to theoretical postulates (1962). Instead, he implied that theories provide the framework for observations, thus controlling the theory that “emerges” from the data. Karl Popper who maintained that observation is theory impregnated (1956/1983) supported this. “Once your eyes were thus opened, you saw confirming instances everywhere: the world was full of verifications of the theory” (Popper, 1980, p. 21). Kuhn and Popper developed the idea that our interpretations of reality based on observations are fallible.

A person's theory of truth or knowledge (epistemic style) is a critical element of one's personal theoretical orientation (Royce, 1964a, Moser, Mulder, & Trout, 1998). Consequently if theoretical orientations are determinants of evaluation practice and outcomes (Adelman, 1996; Everitt, 1996; Pawson, 1996; Scriven, 1996; Swandt, 1997a), epistemic style is also a component.

Epistemic Style and Evaluation

Evidence that personal epistemic style influences the validity of evaluation outcomes directs us to conduct and report evaluation studies in a way that allows the audience of the evaluation report to understand the assumptions under which the evaluation was taken, the barriers to achieving objectivity, and the rationale taken for each evaluation and measurement decision (e.g., why was this particular sample

chosen?) (Altheide & Johnson, 1994; LeCompte & Goetz, 1982; Lincoln , 1990; Lincoln & Guba, 1985; Maxwell, 1992). This allows clients and stakeholders to make their own determination of the significance and meaning of the evaluation report. It also informs the utilization of the evaluation data. In addition, it makes it possible for the reader to combine the evaluation report with other formative evaluation data collected by the organization in order to develop a more holistic picture of the organization, integrating different perspectives.

This study examines the cogency of these statements. Although it is recognized that there are many influences on the evaluation process, this study will focus on the way that epistemic styles of the evaluator shape evaluation practice and the way that meaning is imposed on observations. It examines whether particular epistemic styles fall into categories of evaluation practice. From the literature, evidence can be found for a typology of at least seven categories of evaluation practice: (1) accreditation and accountability (Affholter, 1994; Vedung, 1997), (2) connoisseurial assessment (Eisner, 1991; Vedung, 1997), (3) emancipatory (Lather, 1991), (4) management (Averch, 1994; Vedung, 1997), (5) naturalistic (Lincoln & Guba, 1985; Patton, 1980; Chen, 1990; Rog, 1994), (6) theory-driven (Chen, 1990; Weiss, 1972, 1998) and (7) participatory (Patton, 1980; Alkin, 1990). The central research questions will be:

1. Do evaluators representing different epistemic styles exhibit differences in evaluation practice?
2. Does the evaluator's epistemic style provide a foundation for evaluation practice?

Rationale for the Study

Epistemic Style as a Foundation for Evaluation Practice

As previously stated, there are many influences on evaluation practice. Epistemic style is chosen for this study under the premise that evaluation practice is derived “from different epistemologies and shaped by quite different attitudes, values, goals, assumptions, principles, and beliefs” (Miller, Nelson, & Moore, 1998, p. 379). This stance assumes that epistemic style is the foundation upon which evaluation practice is built. Additional factors build upon this foundation to shape practice. Theoretical orientations are based on what we believe but epistemic style describes how we come to believe. Thus, I believe that epistemic style is not only the foundation for evaluation practice but also the foundation for other theoretical orientations.

Theory-Laden Observations

“Mr. Turnbull had predicted evil consequences...and now doing the best in his power to bring about the verification of his own prophecies.” This quotation from an Anthony Trollope novel and cited by Karl Popper (1980, p.19) is an example of how observations can become theory-laden. The dangers of this happening is that anomalies (observations that do not fit with our theoretical orientation) can be viewed as inconsequential. In the field of program evaluation, anomalies may be viewed as non-compliance, resulting in negative judgments. In “actuality,” anomalies may be important indicators of program quality that could be used to improve programming. In the case of schools, if we evaluate programs only by established criteria, we may be missing important elements that could bring about “school reform.” If this is the

case, observations can have no epistemic significance. In the end, it leads to “perspectivism” as described by Nietzsche, “...there is no sensible conception of a world independent of human interpretation and to which interpretations would correspond if they were to constitute knowledge” (Dancy & Sosa, 1992, p. 304).

Consciousness of Personal Theories

Social constructivism holds that many “truths” we study are socially constructed from our interpretations (Gergen, 1985). This points to the critical importance of researchers being self-reflective, of being conscious of their theory in use. Courses in program evaluation emphasize the idea that theories are grounded in the data. But choices of questions, methods of collecting data, methods of coding and analyzing data, are all developed within a theoretical framework (Fine, 1994; Gubrium & Silverman, 1989; Hamilton, 1994). This framework influences our interpretations of observations and epistemic style will determine the credence that evaluators give particular pieces of evaluation evidence (Huberman & Miles, 1994; Royce, 1964b).

Lack of Empirical Evidence

The way that individual characteristics of the researcher shape the way that theoretical meaning is imposed on experience is not well documented in the literature (Daly, 1997). A survey of the literature also reveals there are few empirical studies substantiating the claim that epistemic style is linked to evaluation practice. It is the intent of this study to provide empirical evidence of the relationship between theory and practice. It examines epistemic style and its relationship to evaluation practice. It is hoped that the study will contribute to an increased awareness of the evaluator

“as an instrument” through which data is filtered.

Demystifying of Theoretical Orientations

This awareness of the researcher as an instrument calls for the writing of evaluation studies in such a way that theoretical orientations (which rest on epistemic style) are apparent in all steps of the evaluation/research process. Research in the post-modern paradigm calls for this demystifying of objectivity in order that an increased awareness of values through which interpretations are filtered is developed (Daly, 1997). Unfortunately, these theoretical orientations and particularly epistemic styles are not always apparent to the evaluators themselves (Banks, 1993; Lather, 1993). This suggests that there is an absence across research and evaluation paradigms of dialogue about epistemological assumptions. Consequently, this may not always be accessible information. Kerry Daly, Associate Professor at the University of Guelph, Ontario, warned us of the dangers of such unawareness.

Post modernity would suggest that we can no longer go merrily along presenting theory as an objective account as if it represented the reality of the participants. However we continue to do this when we do not account for the theorist's self as a thinking, imposing, reflecting, and participating member of the social reality being presented. (1997, p. 349)

Because we often cannot determine the epistemological foundation of an evaluator, I believe the validity of evaluation practices and outcomes are also often in question. This should not be reason for despair, but reason for making the assumption that there is an influence of epistemic style on evaluation practice. This biasing factor can be counterbalanced by incorporating various perspectives into the evaluation process. This study provides empirical evidence for the need for program participants to take part in the evaluation process in order that multiple perspectives are taken into account.

Personal Framework of the Study

This study is built on the theory that personal qualities of the evaluator (such as epistemic style, personality, gender, and biography) have an effect on evaluation practice. For this reason, it is important for the readers of this study to understand my personal story and beliefs. This segment of the study can be found in the Appendix A. The inclusion of autobiography in evaluation and research studies has received ever-increasing attention. Michelle Fine argued that researchers “carry no voice, body, race, class, or gender and no interests into their texts” (1994, p. 74). She denounced this practice and supported her position with a quote from Henry Giroux. “As researchers, we need to position ourselves as no longer transparent, but as classed, gendered, raced, and sexual subjects who construct our own locations, narrate these locations, and negotiate our stances with relations of domination” (Fine, 1994, p. 76). In this study, I will attempt to make my own assumptions and stance apparent to the reader.

Organization of the Study

Chapter I presented an overview of the need to focus on epistemic styles of evaluators. The problem was outlined and a rationale for the study developed. The purpose of the study is to examine the relationship between epistemic style and evaluation practice.

Chapter II presents a review of literature that is pertinent to the research questions. This body of literature is composed of theoretical works that provide the foundation for the conceptual foundation of this study and empirical studies that provide evidence for the substantiation of the premises of this project.

Chapter III describes the methodology used in the project. The rationale for

each component of the methodology is justified.

In Chapter IV, the data are presented and analyzed.

Chapter V summarizes the findings and implications for evaluation practice.

Suggestions for further study in this area and limitations of the study are described.

CHAPTER II

LITERATURE REVIEW

Organization of the Review

Literature examined in this study provides evidence for: (a) influences on evaluation practice and outcomes, (b) a definition of theory and epistemic style (c) influence of epistemic style on behavior, (d) the relationship between epistemic style and evaluation practice, (e) the relationship between epistemic style and evaluation outcomes, (f) typologies of evaluation practice, and (g) literature that develops the rationale for this study.

Influences on Evaluation Practice

Intervening Variables

A great deal has been written about the influences on research methodologies. Less has been written specifically about the context of evaluation practice but as research methodologies are used in the process of evaluation, the same influences predominate. In a study that examines the influence of epistemic style on evaluation practice, it is important to study other influences on evaluation practice in order that possible intervening variables are understood.

Theoretical Orientations

In this study, it is essential that the reader understand the relationship between theoretical orientation and epistemic styles. Epistemic styles (epistemologies) are a

type of theoretical orientation. One's epistemic style is one's theory of knowledge and truth, thus it is a theoretical orientation. I have chosen it for study because I believe it to be central to all other theories held by individuals; it is the thread that runs through all theories. For example, theories of how to best teach children rest on our views of how we believe knowledge is obtained and what constitutes knowledge (Dewey, 1938/1970). Our theories of religion and religious practice also rest on our theories of what is truth and how truth is obtained (Popkin & Stroll, 1986).

Researchers' theoretical orientations begin the process of bounding and framing the research (Marshall & Rossman, 1995). This calls for theoretical candor that Thomas Swandt explained in the following quote:

Similar procedures for documenting both the a priori function of theory—its role in shaping hypotheses, problems, and propositions to be examined as well as the events, objects, interactions, and people the fieldworker looks for in the field. The use of theory in developing local accounts of significance as the fieldwork unfolds is emphasized by many methodologists in qualitative inquiry. (Swandt, 1997b, p. 154)

He maintained that one enters the field with a theoretical language and attitude. This is a major tenet of qualitative evaluation. Jennifer Greene, of Cornell University, stated that qualitative evaluators have "endeavored to make explicit the value dimensions of our work, understood that methodological decisions are also value statements, and struggled to develop defensible subjectivities that help substantiate our knowledge claims" (1996, p. 277). Greene described a doctoral dissertation at Cornell University in which an evaluation was done of a youth program in an evangelical organization (Greene, 1996). In this study, the evaluator explicated his own theoretical orientation, based on his own biography and experiences, that interjected into the evaluation project. By doing this, the evaluator provided a critical framework from which to examine the evaluation findings, thereby

acknowledging that theoretical orientations influence evaluation practice. In the process of evaluation, theoretical orientations will provide the foundation for the design of the evaluation plan (Guba & Lincoln, 1989). It is often the case that the evaluation questions are developed in collaboration with the clients, but the design of the evaluation is the responsibility of the evaluators in most cases. Alasuutari (1996) studied theorizing in qualitative research and emphasized the importance of the researcher reconstructing the framework implied in the questions asked and the research design in question.

Such a more or less explicit theoretical framework consists of ontological and epistemological premises, that is, of notions about the nature of the reality being studied and the ways by which one can study that reality. The main function of data collection and analysis is to make one's own underlying premises as visible as possible and to challenge and develop the initial framework. (Alasuutari, 1996, p. 373)

Political Context

In this study, political context refers to the power structure of the organization or community in which the evaluation takes place. It is typical that those who have power will set the stage for practice within an organization.

Miller, Nelson, and Moore conducted a study that examined difficulties researchers had in developing qualitative research inquiries (1998). The researchers maintained that the political contexts of their universities often “undermined principles of systematic interpretivist design” (p. 381). The researchers, in this study, believed they had to compromise their original designs in order not to jeopardize their funding, degrees, tenure, or promotion. This was particularly true of untenured faculty and doctoral students.

The political context may also be within the evaluation organization. There is generally a hegemonic structure within the group that provides the framework for

both design and reporting of evaluation findings. This can eventuate into political context becoming the primary influence of evaluation design within a group. Miller, Nelson, and Moore (1998) cautioned that if there is not conscious awareness by the researcher of conflicted research texts, “researchers internalize conflicting epistemologies which in turn produce conflicted research voices” (p. 399). Likewise, this would seem to apply to the evaluator. If an evaluator’s “voice” is hidden within the framework of the voice of his/her superiors, the framework of practice becomes a mindset from which it is difficult to break free.

Culture of Organization and Community

Culture of the Evaluation Agency

Often times the culture of the evaluation agency defines what types of evaluation plans, methods, and style are acceptable. Sometimes evaluators are caught in this identity or caught in an identity that is defined by their expertise. If that expertise is an avenue to contracts and further work, they are invested in maintaining that identity of self and the identity of the organization (Bloom, 1998).

Culture of the Community

The culture of the community also can determine evaluation practice. The way in which a community accepts, welcomes, tolerates, or rejects the evaluator influences the ability of the evaluator to implement intended evaluation practices (Guba & Lincoln, 1989). These attitudes are often embedded in the core values of the local culture. It is critical that the evaluators become established at the evaluation site in early stages of the evaluation. Negotiating a relationship with a gatekeeper who can facilitate access to a community (Guba & Lincoln, 1981) can effectively do this.

Time and Budget

Large determinants of evaluation practice in individual evaluations are often the amount of available time and financial resources (Chen, 1990; Worthen & Sanders, 1987). Budgets sometimes do not allow for the comprehensive approach that evaluators or clients wish to conduct. Clients may not be able to afford a lengthy process of evaluation. Surveys and document reviews do not take nearly as much time as interviews, observations, and focus groups; consequently they are often times much cheaper. Thus, budget can determine evaluation practice. Carol Weiss addressed these types of problems in her book, Evaluation (Weiss, 1972/1998).

Time can also be tied into budget concerns. In public-service agencies funding often doesn't allow for as comprehensive an evaluation as desired if the evaluation is being conducted by an external evaluator (Weiss, 1972/1998). If an evaluation agency specializes in particular type of clients, it may appear as the evaluation agency specializes in a particular pattern of evaluation practice when reality is that the type of evaluation methods used are determined by financial resources of clients. There is a current trend in evaluation circles to circumvent this situation by practicing more participatory forms of evaluation such as empowerment evaluation (Fetterman, Kaftarian, & Wandersman, 1996), utilization-focused evaluation (Patton, 1997) and cluster-evaluation (Barley & Jenness, 1995). These evaluation methods draw the clients and stakeholders into the process of evaluation that oftentimes lowers the cost while at the same time builds the evaluation capacity of the organization. In this way, formative evaluation practices can be embedded into the on-going activities of the organization.

Gender and Race

Gender is a major influence on the development of research and evaluation practice. So much so, that a category of research and evaluation practice identified as feminist has been well developed.

Patti Lather, a professor of Educational Policy and Leadership at Ohio State University, defined feminist research as research that puts the social construction of gender at the center of inquiry (1991). It is an emancipatory stance that has at its goal to “correct both the invisibility and distortion of female experience in ways relevant to ending women’s unequal social position” (Lather, 1991, p. 71). Lather described how as feminist research practice evolved, methods evolved to a more interactive, contextual process to search for patterns of meaning. Prediction and control lost stature for this group of researchers in favor of a dialectical process in which both researcher and researched negotiated and explored meanings.

This dialectical process was a foundation of the design of a research project in California of Latina women. Esther Madriz, a professor of sociology at the University of San Francisco, conducted a study of lower socioeconomic status Latina women on the topic of fear of crime (Madriz, 1998). Using focus groups within a feminist methodology framework, Madriz attempted to remove the barriers of communication between Latina women and traditional forms of research. It was critical to her intent to diminish the distance between the researcher and the researched in order to understand the story of the Latina women. By the use of focus groups, Madriz created a communication process (collective testimonies) that was familiar to the women. Madriz explained her rationale for the use of focus groups: “The interaction in focus groups emphasizes empathy and commonality of experiences and fosters self-disclosure and self-validation” (Madriz, 1998, p.116).

Certain facets of the Latina women's lives though were problematic to conducting the research. In traditional Latino culture, women were solely responsible for their children and their house. This made it difficult for them to arrange to leave. Often there were complications at the last moment. She also cited the problem of people being on time, a problem she partially attributed to the Latino culture. Madriz attempted procedural remedies for these problems. Sometimes she arranged rides or payment for participants. She also scheduled for women to arrive a half-hour before the focus group was to begin. The fact that she was a Latina woman also engendered trust within the group. This study provided an example of how research practice is influenced by gender (using a method comfortable for women) and by race (making accommodations that facilitated participation).

A prime example of a body of work influenced by race, gender, and class is the work of bell hooks, professor at City College, New York. Her experiences of school in an African-American community "enable hooks to affirm her right as 'subject' in spite of the oppressive forces of racism, sexism, and classism, and to articulate an educational pedagogy that seeks to develop and nurture critical consciousness in both teachers and students" (Florence, 1998, p. xvii). Florence described hooks' work as a re-conceptualization of the knowledge base (1998). Although bell hooks is defined as an educator rather than an evaluator and researcher, her work is valid evidence of the effect of race and gender on practice.

Experience of Evaluator and Client

The experience of the "researched" or the client with prior research projects or evaluations can profoundly affect their receptivity to research and evaluation practice. Leslie Bloom found this to be a problem in her study of a woman school

administrator (1998). She quoted from a doctoral dissertation by Petra Munro at the University of Oregon to illustrate the problem.

I sensed that my request was perceived as a demand, which did not conform to my participants' conceptualization of the research process. My heightened sensitivity to avoiding an exploitive research relationship had not taken into account the fact that my participants had their own reasons and agendas for participating in the study. In essence, my assumption of the need for a collaborative relationship underscored my perception of them as dis-empowered, thereby disregarding their power to determine the nature of the relationship. (cited in Bloom, 1998, p. 45)

Relationship of Evaluator and Client

The relationship between an evaluator and client can often determine what types of evaluation activities can take place. This has been well documented in the field of research, particularly ethnography and feminist methodologies (Bloom, 1998). Trust between the evaluator and the evaluatee is generally a component as well as the evaluatee's willingness to participate in the study/evaluation. This willingness of course may be a factor of the trust component.

In a study conducted in New Jersey in 1993 (Wong, 1998), power relationships facilitated respondents enthusiasm to participate in the study. This study examined the effect of the New Jersey Welfare Reform Act of 1991 and how it affected the lives of women and children on Aid to Dependent Children (ADC). The respondents believed that the researcher would be able (as a psychologist and an authority figure) to help them with their personal problems. Of course this was not part of the purpose for the interviews and focus groups. Nevertheless, the researcher exploited the respondents' need for counsel to gain their participation in the research project. In retrospect, the researcher realized that the dynamics of the interviews and focus groups were likely influenced by the power relationships that existed, and by

the fact the women were financially needy and were paid for their participation. The women's need for psychological help, their need for a sympathetic ear, and a hegemonic culture of deferring to authority in order to obtain necessities created a fertile ground for this particular researcher. Michelle Fine quoted Joyce Ladner to explain this dilemma:

It has been argued that the relationship between the researcher and his subjects, by definition, resembles that of the oppressor and the oppressed, because it is the oppressor who defines the problem, the nature of the researcher and to some extent the quality of interaction between him and his subjects. (Fine, 1994, p.73)

Definitions of Theory and Epistemic Style

Definition of Theory

Epistemic style, which is the foundation of this study, is a theory. This section will clarify the definitions of theory and epistemic style that I use in this study. A dictionary defines theory as "a belief, policy, or procedure proposed or followed as the basis of action" (Merriam-Webster's Collegiate Dictionary, 1997, p. 1223). Shaddish, Cook, & Leviton (1991), defined theory as a body of knowledge that organizes, categorizes, describes, predicts, explains, and otherwise aids in understanding. This study takes the view that theories are collections of models upon which action is based. The question becomes whether observations are determined through the lenses of these theories (models). If we agree with the premise that observations are framed by our theories, it is important to know how these theories are formed (Scriven, 1996). Are they based solely on observations or do they spring from a priori, non-inferential beliefs?

Definition of Epistemic Style

Epistemology is “the study or a theory of the nature and grounds of knowledge especially with reference to its limits and validity” (Merriam-Webster’s Collegiate Dictionary, 1997, p. 390). In his study of individuals’ personal epistemologies, Joseph Royce established a typology of epistemology that he called epistemic styles: rationalism, intuitionism, empiricism, and authoritarianism (1964). Each of these styles has a particular criterion for truth. These criteria give us the procedure for telling us what is true or false. These styles are defined in the following paragraph.

Rationalism states that nothing is true if it is illogical. This perspective demands that thought be logical. If thinking is fuzzy, truth has not been found .

Empiricism demands observation. This is interpreted as the need to touch, see, or hear something before it can exist. This obviously limits truth to our experiences or access to the experiences of others. As an example, an empiricist would accept the research findings conducted by others as long as the findings were based on sense perceptions.

Authoritarianism assigns truth on the basis of authority. Authority may be derived from position (school principal, professional evaluator, church minister, and our parents) or may be a honored document (the Bible, the textbook, an encyclopedia, a biography).

The last epistemic style described by Royce was intuitionism that claims truth by direct apprehension. Royce quoted Goethe to describe this perspective.

It is not given to us to grasp the truth, which is identical with the divine, directly. We perceive it only in reflection, in example and symbol, in singular and related appearances. It meets us as a kind of life that is incomprehensible to us, and yet we cannot free ourselves from the desire to comprehend it. (Royce, 1964, p. 129)

Theories of Knowledge

The literature establishes there are many theories of knowledge upon which various disciplines build conceptions of truth (Dancy & Sosa, 1992; Moser, Molder, & Trout, 1998).

Foundationalism

Foundationalism is built on the premise that there are basic beliefs, which are justified independently of any other belief. A strong foundationalist maintains that these are infallible and incorrigible (Lewis, 1952; Bonjour, 1985). They are what we believe without reference to any other beliefs; they stand alone. This does not seem to allow any room for observations to be theory-laden. It implies that our observations must be based on a foundation of truth and this foundation needs to be incorrigible and known without experiential claims being used as premises.

Conjectures and Refutations

Karl Popper, philosopher and Professor of Logic and Scientific Method at the London School of Economics, adamantly opposed the idea of infallibility. He proposed that knowledge be developed through a process of conjectures and refutations (Popper, 1956/1983). He wrote “our theories, even the most important ones, and even those which are actually true, always remain guesses or conjectures. If they are true in fact, we cannot know this fact; neither from experience, nor from any other source” (1956/1983, p. 33).

Coherentism

Susan Haack, a contemporary philosopher of epistemology, attempts to solve

this problem in her book Evidence and Inquiry (1993) with a theory of coherentism. From her perspective, beliefs are justified through a combination of experiential evidence (observations) and foundational beliefs. In her view, it is the balance of observations with foundational beliefs that makes justification evidence truth indicative.

Social Constructivism

Social constructivism holds that many “truths” we study are socially constructed from our interpretations (Gergen, 1985). It is a perspective that studies the way people construct and interpret (both individually and collectively) their world in specific contexts which may be social, historical, or psychological (Swandt, 1997b). This seems to be similar to Swandt’s definition of cognitive relativism that there are no universal truths (1997b).

This sample of theories of knowledge does not preclude other theories, it only establishes there are many theories upon which we base our behavior.

Epistemic Styles and Behavior

There is an empirical study in the literature that links epistemic styles to behavior. In 1989, a study was published that examined whether epistemic style was linked to the selection of counseling approaches (Neimeyer, Prichard, Lyddon, & Sherrard, 1989). The Psycho-Epistemological Profile was used and the results were correlated with participants’ selection of a description of counseling practice. Researchers found the strongest correlation between rational styles and a rationalist approach ($r = .17$), metaphoric styles and social constructivist approaches ($r = .35$), and empirical styles and behavioral approaches to counseling ($r = .28$).

Although empirical studies examining epistemic style can be found, there is little evidence that studies the link to behavior. This points to the need for such empirical studies to be conducted.

Epistemic Styles and Evaluation Practice

Several researchers make a distinction between the terms methods and methodologies (Bloom, 1998; Greene & Caracelli, 1997; Harding, 1987; Stanley & Wise, 1990; Taylor & Rupp, 1991). As Leslie Bloom explained, “methods are research techniques, procedures, and practices, and methodologies are the theories or perspectives that inform the production of particular kinds of research and justify it in terms of its knowledge making” (Bloom, 1998, p. 138).

Thomas Kuhn, in his book, The Structure of Scientific Revolutions (1962), maintained that operations and measurements are paradigm determined. The scientist functions within the paradigm of “normal science” that falls under the old adage “what you see is what you get.” Particular instruments allow us to see only particular things. Selection of method and instrument determines what we can observe and it is from this premise that Kuhn believed that data are not unequivocally stable (Kuhn, 1962). To frame the question differently, if scientists, holding different theories of what they expect to see, look through the same microscope, do they see different things even though the slide doesn’t change? Assuming normal eyesight, they all see the same thing. *How* the slide is interpreted of course may differ among scientists. It is this point that is not arguable as interpretations are considered to be fallible (Popper, 1956/1983).

Epistemic Styles and Evaluation Outcomes

Standardization as Coercion

There are claims in the literature that evaluators' theories of knowledge (epistemic styles) will have an effect on evaluation outcomes (Denzin, 1994; Morse, 1994). Epistemic style can determine what sources of evidence will be given credence, and what evidence actually is sufficient to make a determination of value or to make a decision about programming or policy. Traditional positivist research relies on the rigor of methods such as random sampling to produce valid knowledge (Creswell, 1994; Dancy & Sosa, 1992). An evaluator using such methods would provide authoritative evidence perhaps without considering alternative explanations (Denzin, 1994). Timothy McGettigan of Wake Forest University, built the argument that "all attempts at standardization—even, or perhaps especially, if they are geared toward an evocation of truth are essentially efforts to impose coercive, ideological constraints on knowledge" (1997, p. 367). In his study of a Green Tortoise adventure travel tour, he showed how during the midst of the tour, a personal "epiphany" changed his framework for the study, thus altering outcomes. He made use of "metaphor" to communicate understanding. Researchers within the postmodern paradigm argue that it is not possible to conduct research without subjectivity and actively search for alternative explanations (Helshusius & Ballard, 1996; Miles & Huberman, 1994; Phillips, 1990). It is reasonable to expect that an evaluator working in a postmodern paradigm would also experience "epiphanies" that translate the design or findings of the evaluation. In both of these situations, epiphany can be defined as a sudden manifestation or perception of the essential nature or meaning of something or an intuitive grasp of reality through something (as an event) (Merriam-

Webster's Collegiate Dictionary, 1997). Evaluators working within the positivist paradigm will likely be constrained by their methodology and will have to either dismiss the "epiphanies" as irrelevant to the project or redesign the entire study.

Evaluation Conveying New Understandings

Patti Lather, noted feminist researcher, strives to create representations that push those that are the subjects of the research and those that read the research to new meanings that are not viewed as final elements of knowledge but new understandings (1993). In a 1993 article published by the *Sociological Quarterly*, she described the difference in how outcomes are determined by looking at the differences in how researchers determine validity. The foundation for validity is based on a theory of knowledge (epistemology). She delineated this argument by describing four ways of ascertaining the validity of the research: ironic validity, paralogical validity, rhizomatic validity, and voluptuous validity. For each she cited an empirical study as an example. Two of these, ironic validity and rhizomatic validity, are described further in this review to provide support for epistemic styles influencing evaluation outcomes.

Ironic validity provides evidence that there are problems with representation by showing there can be more than one interpretation of research. Lather used as an example a study conducted in 1941 by James Agee and Walker Evans (1993). Their text, Let Us Now Praise Famous Men, examined the devastation of rural America during the Depression. The text begins with Evan's uncaptioned photographs and ends with Agee's text. Agee included multiple endings and also reflected how his subjectivities and autobiography influenced his interpretation. Although their study was not a traditional evaluation project, it aptly illustrated that there are multiple

outcomes depending on subjectivities of the researchers and those who are the consumers of the research. The photographic essay at the beginning allowed for dynamic outcomes across time rather than a static finality. The choice in form of representation by these men, grew out of what they considered evidence for truth. In turn, their style of representation allowed for multiple outcomes. It appears from the study that Evans and Agee did not subscribe certain truth and consequently did not bring closure to their study.

Rhizomatic validity, as described by Lather (1993), looks for networks of questions that develop within the research. Where are there new possibilities? Researcher assumptions are interrupted. Tensions between researched and researcher remains within the text in order that interstices can be explored. Lather maintained that Derridean rigour is an important aspect of this type of validity. Based on the theories of Jacques Derrida, this means that the research (or evaluation) attempts to undermine the center to show what is marginalized as well as to show multiple meanings and possibilities.

Postmodern theorists provide evidence that personal theories have an influence on the outcomes of research. In the field of evaluation, this is a critical issue as policy development often relies on evaluation reports (Worthen & Sanders, 1987). Evaluation is not strictly a technical field of practice using a set of methods. Consequently, it seems important to understand the theories different categories of practice rest on. The following section will describe a typology of evaluation practice.

Typologies of Evaluation Practice

Categories of Practice

Categories of evaluation practice change as the paradigms of social science research evolve and modulate with the dynamic nature of society. In this study, evidence was found in the literature (drawing heavily from Worthen and Sanders' 1987 book, Educational Evaluation) for six categories of evaluation practice. These categories are not discrete, but have components that overlap with each other. The following categories are based on practices that are developed around purposes of evaluation. This study will later examine whether a typology can be developed around epistemic styles of evaluators.

Accreditation and Accountability

Evaluation for accreditation and accountability purposes is common to government agencies such as schools, health agencies, and welfare agencies. In this framework, it could be considered a continuous monitoring feature of public decision-making systems (Vedung, 1997). Evaluations conducted for accountability reasons determine whether an agency is doing its job. Accreditation evaluation examines whether an agency has met minimum standards (Vedung, 1997). Both could be characterized as a "policing" action and both are based on measuring achievement of established objectives. Worthen and Sanders (1987) described accreditation approaches to evaluation: (a) published standards, (b) a self-study by the institution (c) a team of external assessors, (d) a site visit, (e) a site-team report on the institution, usually including recommendations, (f) a review of the report by some distinguished panel, and (g) a final report and accreditation decision by the

accrediting body.

In the present accountability environment, this type of evaluation has received prominent attention. In 1993, the U. S. Congress passed the Government Performance and Results Act of 1993 (GPRA) to provide for strategic planning and performance measurement in the Federal Government (<http://www.financenet.gov/financenet/fed/cfo/gpra/gpralaw.htm>). The purposes were to: (a) improve the confidence of the American people in the capability of the Federal Government by systematically holding Federal agencies accountable for achieving program results; (b) initiate program performance reform with a series of pilot projects in setting program goals, measuring program performance against those goals, and reporting publicly on their progress; (c) improve Federal program effectiveness and public accountability by promoting a new focus on results, service quality, and customer satisfaction; (d) help Federal managers improve service delivery, by requiring that they plan for meeting program objectives and by providing them with information about program results and service quality; (e) improve congressional decision-making by providing more objective information on achieving statutory objectives, and on the relative effectiveness and efficiency of Federal programs and spending; and (f) improve internal management by the Federal Government. This Act has significant effect on the evaluation community. It means that all programs and agencies receiving funding from the Federal government must have evaluation components framed by GPRA standards. In many cases, the requirement for comprehensive accountability oriented evaluations precludes the possibility of doing more emancipatory or participatory evaluations.

GPRA has focused attention on objective-oriented evaluation. The focus is on whether measurable objectives have been achieved. The evaluators' responsibilities

include the development of objective instruments to gather data that is reliable and valid. The task becomes looking for discrepancies between objectives and performance.

Emancipation

In the preface to Getting Smart, Michael Apple, described Lather's research as having an empowering approach to generating knowledge (Lather, 1991). This is an apt description of emancipatory evaluation. Apple wrote:

Coming close to Paulo Freire, Lather proposes a new, more emancipatory way of validating critical research, what she calls *catalytic validity*. Rather than researcher neutrality, she argues for a more collaborative, praxis oriented and advocacy model that acts on the desire for people to gain self-understanding and self-determination both in research and in their daily lives.... The norms guiding such work involve open-endedness, dialogue, explicitness, and reciprocity. (Lather, 1991, p. x)

Although there are similarities between this type of evaluation and participatory models, it rests heavily on critical theory and the feminist model of research. Its goal is to enable people to recognize the inequities and to build capacities for equitable change in communities and agencies. The end product is considered to be justice.

Expertise

Often times, a recognized expert in a field is called to provide value judgments of a program or a product. Such professional judgments of quality are called expertise-oriented evaluations (Worthen & Sanders, 1987).

Elliot Eisner, professor at Stanford, described connoisseurship as the art of appreciation and criticism as the art of disclosure (1991). In his book, The Enlightened

Eye (1991), Eisner maintained that educational criticism is an evaluation with an educative purpose conducted by recognized experts in education. This positions the education critic or connoisseur as someone who attempts to enlarge perceptions and deepen understandings. In Eisner's viewpoint, truth can never be found with certitude but always involves interpretations and judgments. This is in sharp contrast to most expert-oriented evaluations, which bear more similarities to the accreditation approach.

Generally, expert oriented evaluation rests heavily on established criteria and guidelines about which the expert is knowledgeable (Worthen & Sanders, 1987). The evaluator is considered the authority. The evaluatees have little input into the process of evaluation. A mixture of qualitative and quantitative methods of data gathering is used in this model of practice.

Management

This approach to evaluation is oriented towards decision-making and focuses on who will use the evaluation results and how they will be used (Worthen & Sanders, 1987). It is generally a formative evaluation practice in which evaluation activities take place during the implementation of a program in order that evaluation results can serve decision-making by management. Objectives do not frame the evaluation questions. Rather, the questions are determined by what type of data is needed to make particular decisions. For this reason, decision-makers (managers and other administrators) take an active part in the evaluation design stage.

Daniel Stufflebeam of Western Michigan University developed a popular management-oriented model for evaluation. His evaluation framework (known as CIPP) is based on four different types of educational decisions (Stufflebeam, 1983):

1. Context evaluation is designed to serve planning decisions. This involves determining the needs in the particular context.

2. Input evaluation is designed to serve structuring decisions. This involves determining what resources are available, what alternative strategies for the program should be considered, and what plan seems to have the best potential for meeting needs.

3. Process evaluation is designed to serve implementing decisions. This step judges how well a plan is being implemented. It examines barriers to success and possible revisions.

4. Product evaluation is designed to serve recycling decisions. This involves measuring outcomes.

The CIPP model follows a very structured process that defines criteria for each decision situation by specifying variables for measurement and standards for use in the judgment of alternatives (Worthen & Sanders, 1987). But management-oriented approaches are not necessarily objectives-oriented or quantitative. An example is utilization-focused models as described by Michael Patton (1997). Lee Cronbach maintained that qualitative methods are helpful in making findings more useful to stakeholders (1982). Cronbach also emphasized integrating program evaluation into the environment of political and organizational processes to facilitate decision making that includes multiple perspectives. Besides their decision-making purpose, the management-oriented evaluations are characterized by the authority (expertise) of the evaluator and by being management-driven.

Naturalistic and Participatory

Worthen and Sanders cite Robert Stake as the first evaluator in the 1960's to

emphasize the importance of including stakeholders into the evaluation process (1987). They define the characteristics of this approach as (a) depending on inductive reasoning, (b) using a multiplicity of data assimilated from many sources, (c) not following standard evaluation plans, and (d) the recording of multiple realities.

Stake evolved from his countenance theory of evaluation that emphasized the two countenances of description and judgment to his theory of responsive evaluation that focused on the concerns and issues of stakeholders in the evaluation (Worthen & Sanders, 1987). Guba and Lincoln echoed this model in their book Effective Evaluation (1981), in which they hold the different value perspectives of stakeholders as paramount. Naturalistic evaluation uses naturalistic inquiry methods as delineated by Guba and Lincoln (1985). This model holds that given enough immersion in the field, truths will be recognizable but multiple.

In the naturalistic approach, sensitive observers record behavior in “natural settings,” and researchers analyze the resulting protocols with due regard for the humanity of the subjects and the desires of program operators, administrators, and other stakeholders. The full complexity of human behavior is thus given due attention, free of the constraints imposed by the research designs of the positivistic approaches. (Chen, 1990, p.21)

Naturalistic models of evaluation involve both qualitative and quantitative methods of data gathering but qualitative methods almost exclusively dominate evaluation designs.

Other participant oriented approaches include David Fetterman’s Empowerment Evaluation (1996) and Michael Patton’s Utilization-Focused Evaluation (1997). Empowerment Evaluation is the use of evaluation concepts, techniques, and findings to foster program improvement and build capacities of self-determination in its stakeholders (Fetterman, 1996). Fetterman described it as focusing on empowering processes and outcomes by the intense involvement of

stakeholders in the evaluation process. It is a democratic process facilitated by the evaluator but directed by the stakeholders. Criteria are determined by everyone. The authoritative voice becomes a chorus of all involved. Utilization-Focused Evaluation, as its title intimates, has use of evaluation findings for program decision-making and improvement as its primary goal (Patton, 1997). In order that evaluations are designed specifically to meet the needs of stakeholders (which include both program participants and decision-makers), stakeholders are included in the evaluation process from the design stage through the data analysis. This process involves building the organization's evaluation capacity. As evaluation becomes part of the organization's embedded practice, the goal is that evaluation findings will also become a critical element of the organization in terms of decision-making and the improvement of programs. Both Fetterman's and Patton's approaches use a variety of qualitative and quantitative methods. Their defining characteristic is a heavy involvement of stakeholders in determining evaluation questions, criteria, and judgments as well as making stakeholders responsible for effective use of evaluation findings.

Theory-Driven

Huey-Tsyh Chen thoroughly developed the rationale for theory-driven evaluations in his book, Theory-Driven Evaluations, (1990) and described the model comprehensively as well. In this book, Chen cited a 1985 study that reviewed 175 evaluation studies reported in journals to support his argument that program theory has been long-neglected in program evaluation. This study claimed that most studies "failed to integrate a prior theory into evaluation in terms of formulating program elements, rationale, and causal linkages" (Chen, 1990, p.18).

The major tenet of theory-driven evaluations is that by placing program theory

at the center of the evaluation design, the theory will provide a framework for identifying issues in need of evaluation, in determining appropriate evaluation methods, and informing utilization of evaluation outcomes (Chen, 1990). This type of evaluation examines program theory that explains how program resources are going to affect desired outcomes. This means that the processes of implementation (intermediary steps) are studied as well as long-term outcomes. In this way, it can be tested whether the failure of a program is due to failure of a theory behind a program or failure of intended implementation processes (Bickman, 1987). By articulating a program theory that specifies goals, purposes, and outputs, stakeholders will have more realistic expectations of program outcomes. In the literature, this has been described as a logic model for program implementation and evaluation (Yin, Kaftarian, & Jacobs, 1996; Patton, 1997). This model specifies program objectives, purposes, output, activities, and intended outcomes, along with connecting assumptions. For each of these, quality indicators and means of verification are also specified. Although the program theory drives the evaluation, there is allowance for multiple perspectives, methods of data gathering and analysis, and inclusion of views from multiple perspectives (Cook, 1985).

The preceding review illustrated only six categories of evaluation practice. This study will attempt to provide evidence that there are other categories of evaluation practice to consider. It will then examine whether particular epistemic styles are related to each of these particular categories.

Rationale for Study

Evaluation Use

Donald Schon (1983, 1987) argued that our society tends to separate the

authorities and sources of knowledge from knowledgeable practice. According to Schon, this technical form of knowledge does not provide a sufficient basis for actions. The practice of using an external professional evaluator to evaluate public and private programs is common (House, 1993). According to Schon's theory, this could jeopardize the possibility of strong utilization of evaluation results. Evaluation that takes place during the course of a program may be intended to serve a formative function to improve programming before the program is complete, but actually be giving summative information that is not being fed back into the program. Those who possess technical knowledge (professional evaluators) must be continuously interacting with program participants in order that the evaluation has meaning and possibilities of utilization by the program participants (Patton, 1997). "So much of what is suggested to teachers and school administrators is said independent of context and often by those ignorant of the practices they wish to improve" (Eisner, 1991, p.11). It is critical to involve those stakeholders who understand the assumptions behind the program being evaluated and are knowledgeable about the original intent of the program in order that the evaluation of the program have utility and validity (Patton, 1997). Those who know the tradition, understand the history, are familiar with those genres, and can see what those settings and practice consist of are most likely to have something useful and informed to say about them (Eisner, 1991, p. 3). Public agencies such as schools cannot afford evaluations that are driven by the evaluator's personal theories alone. There is too much at stake.

Observations Are Theory-Laden

Wagner (1993) studied researcher's positions across fields and over time and found that researchers all have blind spots derived from different research dialogues,

processes, and epistemological histories. Consequently, he posited that we need epistemological dialogue to help us understand the gaps in our knowledge and locate our expertise on the broader map of the field.

Pertti Alasuutari, a professor in the Department of Sociology at the University of Tampere in Finland, wrote about the relationship of empirical research to theory in qualitative research.

Being theoretically informed means that one is reflexive toward the deceptively self-evident reality one faces in and through the data, able to toy with different perspectives to it, and that one is open to new insights about everyday life and society. (Alasuutari, 1996, p. 374)

Alasuutari believed that it is critical that a researcher is conscious of his epistemological foundations. In his study of alcoholics in Finland, Alasuutari used a reflexive process to realize that his focus on dart players as a cultural group was shutting out learning about personal motives of dart players on an individual level.

The main task of the researcher is to dig out and reconstruct the framework implied in the questions asked and in the research design in question. Such a more or less explicit theoretical framework consists of ontological and epistemological premises, that is, of notions about the nature of reality being studied and the ways by which one can study that reality. The main function of data collection and analysis is to make one's own underlying premises as visible as possible and to challenge and develop the initial framework. (Alasuutari, 1996, p. 373)

If this critical reflexive process is used during the evaluation process, the evaluator may discover how the addition of different perspectives and other 'types' of evidence give a grander picture.

Awareness of Evaluator as Instrument

An awareness of the evaluator as an instrument through which observations are filtered is akin to a major tenet of poststructuralism—the deconstruction of text to

illuminate distortions and biases. James Banks, in an article on multicultural education, described a study by W. E. B. DuBois conducted in 1935 that made historians aware that their regional and cultural biases were influencing their interpretations of the Reconstruction period (Banks, 1993).

Feminist researchers have integrated this need to identify their ideological positions and assumptions into their studies (Banks, 1993; Bloom 98). It is part of what defines research as feminist and distinguishes the particular paradigm of research.

Multiple-Perspectives

The value of incorporating many perspectives into research and evaluation has been supported in various disciplines from sociology to education to psychology.

Joseph Royce, a professor of theoretical psychology, in his book, The Encapsulated Man, quoted Erwin Schrödinger.

The isolated knowledge obtained by a group of specialists in a narrow field has in itself no value whatsoever, but only in its synthesis with all the rest of knowledge and only inasmuch as it really contributes in this synthesis something toward answering the demand, “Who are we?” (Royce, 1964, p. 1)

Royce’s theory was that man is encapsulated, claiming to have all of the truth when it is only possible to have partial truths. Royce maintained this creates an epistemological dilemma—if truth can always only be partial, how can we ever find meaning? He posited that we can escape this despair by using multi-perspective and multi-disciplinary approaches.

Leslie Bloom, a qualitative researcher from Iowa State University, explored the value of using multiple perspectives within a research study. She studied Feminist Research from within—by doing it and putting herself under the microscope

(Bloom, 1998). She studied two women, a university faculty member and a school administrator using the reciprocal process of feminist methodology. She defined feminist methodology as research in which the researcher does not speak for the respondents, but within a dialogue that integrates multiple perspectives, including the respondents.

Elliot Eisner also built the rationale for using multiple perspectives. Eisner, in his book The Enlightened Eye, wrote, “the selection of a form through which the world is to be represented not only influences what we can say, it also influences what we are likely to experience” (Eisner, 1991, p. 7). He later illustrated this with the phrase, “a way of seeing is also a way of not seeing” (Eisner, 1991, p. 63). Eisner maintained that inquiry in the field of education will “be more complete and informative as we increase the range of ways we describe, interpret, and evaluate the educational world” (Eisner, 1991, p. 7).

Karl Popper advocated an active searching out of multiple perspectives. Popper believed that we must break free from our assumptions and be willing to criticize and refute our theories, before we can have growth in knowledge (1970). The aim of science, according to Popper, is to increase the truth content of our theories by eliminating sources of error. This elimination of error would be by the process of conjectures and refutations. This view points to the need to involve various perspectives in the evaluation process (particularly practitioners). “Rather than wait for errors to reveal themselves, perhaps with disastrous consequences, we consciously and deliberately seek them out: we put our ideas and inventions to the test, we probe critically, we scrap what may be wrong and try again” (Miller, 1985, p.9).

A way to put our ideas and inventions to the test is to expose ourselves, as

evaluators, to different perspectives and to develop an “epistemology of pluralism,” (The New London Group, 1996, p. 13). This does not mean evaluators should just have an understanding of different perspectives of evaluation but to actually seek out and incorporate different and conflicting perspectives in order to enrich the findings of an evaluation (Greene & Caracelli, 1997).

Summary

The literature establishes the idea that there are many influences on evaluation practice. This study examines one particular influence—the personal epistemic style of the evaluator. By this study, I hope to provide empirical evidence of the relationship of epistemic style to evaluation practice. These findings are intended to have implications for the way we design program evaluations and report outcomes.

CHAPTER III

METHODOLOGY

It has been established that theoretical orientations differ among researchers. This study examines the differences between personal epistemic styles of professional evaluators and the influence this may have on evaluation practice.

Differences in Epistemic Styles Among Professional Evaluators

Selection of Instrument

The instrument used to study epistemic style was the Psycho-Epistemological Profile (Royce, Mos, & Learsley, 1975). A literature review revealed few instruments containing measures of epistemology. Other instruments considered for this study were the Study of Values (Allport, Vernon, & Lindzey, 1960), the Theoretical Orientation Scale (Coan, 1979), and the Keirsey Temperament Sorter (Keirsey & Bates, 1978). The Psycho-Epistemological Profile was chosen for its focus on epistemic style and ease of administration and analysis.

Description of Psycho-Epistemological Profile

The Psycho-Epistemological Profile (PEP) was developed at the University of Alberta in the Center for Advanced Study in Theoretical Psychology with the intent that it would provide a profile of an individual's epistemological hierarchy (Royce, Mos, & Kearsley, 1975). The book, The Encapsulated Man (Royce, 1964), provided the conceptual framework for the instrument. Royce identified three basic

ways of knowing: (1) symbolizing, (2) conceptualizing, and (3) perceiving, which he correlates respectively to three epistemic styles: (1) metaphoric, (2) rational, and (3) empirical (Royce & Mos, 1975). The PEP, developed in 1975, measures the ways in which a person combines metaphoric, rational and empirical styles. This results in a Psycho-Epistemological Profile with a score on three scales representing the three epistemic styles. These styles were conceptualized by the authors as higher-order personality factors, which determined individuals' particular worldviews (Royce, Mos, & Kearsley, 1975). Each individual's epistemological hierarchy integrates the three epistemic styles but it is theorized by the authors of the instrument that one style will be dominant and will provide the framework for individual action. In the manual for the PEP the rationale for using the instrument is explained.

The implication of epistemic styles is that individuals with different epistemological commitments possess limited or encapsulated images of reality as a function of their particular epistemological hierarchy. The different epistemic styles represent legitimate approaches to reality, but there are different psychological processes and truth criteria underlying each style. The cognitive processes underlying these three ways of knowing may lead to error. However, each is capable of also leading to truth. Furthermore, it is recognized that these cognitive processes do not function independently of each other. While mankind needs to invoke all the available ways of knowing in order to gain the most comprehensive world-view, individual men tend to be partial to one or the other of the cognitive processes, thereby reflecting different epistemological hierarchies. (Royce, Mos, & Kearsley, 1975, p.2)

If evaluators have different criteria for truth, program evaluations should include multiple perspectives in order that knowledge is not partial or "encapsulated." Table 1 on the following page is based on information in the original manual for the instrument and further describes each of the metaphoric styles.

Table 1
The Basic Paths to Knowledge

Epistemic Style	Cognitive Processes	Truth Criteria for Determining Validity
Rationalism	Conceptualizing. Clear-thinking and the rational analysis and synthesis of ideas	Logical-illogical. Tests validity of views by their logical consistency
Empiricism	Perceiving. Commitment to external experience. Active perception and the seeking out of sensory experience.	Perception-misperception. Tests views of reality in terms of reliability and validity of observations
Metaphorism	Symbolizing. Symbolizing nature, including both conscious and unconscious aspects	Universality-idiosyncratic. Tests validity of views by the universality of his/her insight or awareness

Development of the PEP

Originally the development of the PEP began with the assumption that there were four basic “ways of knowing”: (1) rationalism, (2) empiricism, (3) intuitionism, and (4) authoritarianism. Working with graduate students in an advanced course in psychometric theory, two equivalent forms with 75 items each were developed. The test was piloted on a group of graduate students in mathematics, biology, and the classics. Later, W. A. S. Smith revised the 150 items and created a new 101 item test (each item with four choices). This test, the University of Alberta Epistemological Profile, was administered to fifty students in four groups: (1) literature and fine arts (intuitive), (2) biology and chemistry (empirical), (3) mathematics and theoretical

physics (rational), and (4) career military officers (authoritarian) (Royce, Mos, & Kearsley, 1975). In this profile, the respondents were asked to rank order the choices for each item. After the instrument was piloted and analysis was completed, the authoritarianism dimension was eliminated for three reasons (Royce, Mos, & Kearsley, 1975):

1. Authoritarianism is not an independent way of attaining knowledge, but is a way of treating knowledge once it has been arrived at via another path (empirical, rational, or metaphoric).
2. It was rare during the pilot testing for a subject to have the highest score on the authoritarianism dimension.
3. Components in the authoritarianism dimension did not meet the criteria set for internal consistency.

In 1964, a third revision (Revised Experimental Form III) was established and piloted. This form eliminated the authoritarianism dimension and other items that did not meet internal-consistency criteria. This revision had 81 items with three choices. Piloting of this instrument showed that the dimensions were not interdependent (a score on one dimension affected scores on other dimensions). In the next revision (Revised Experimental Form IV, 1966), items were translated into statements which respondents could disagree or agree. A subsequent version (Revised Experimental Form V, 1968) made minor language changes. In Form V, each of the randomly ordered items on the final forms pertains to one of the three ways of knowing: (1) empirical, (2) rational, or (3) metaphoric. There were 30 items measuring each dimension and the respondent was asked to indicate agreement on a five-point scale.

The sixth revision (Revised Experimental Form VI) changed two items that individuals had difficulty answering or frequently left blank. The numerical scale

was changed to five preference categories: complete disagreement, moderate disagreement, neutral, moderate agreement, and complete agreement. Essentially, Versions V and VI were the same in form and content.

Norming Population of the PEP

The PEP (Revised Experimental Form V) was standardized on a junior college consisting of 925 males and 417 females from ages 19=24. Among this population there were no extreme scores or consistent preference for one epistemic style. There were consistent differences between the means on the three epistemic styles for male and female subjects. The data suggested that these differences could be attributed to males' commitment to empiricism and females' commitment to rationalism (Royce, Mos, & Kearsley, 1975).

Reliability of the PEP

Reliability of the PEP was established through: item analysis, split-half reliability, test-retest reliability, and intercorrelations among the dimensions. These studies were conducted by the developers of the test.

The item analysis yielded positive correlations for each of the 30 items under each dimension with the total score.

The split-half correlational study correlated odd and even items from an administration of Form V to 142 first-year university students and of Form VI to 95 first-year university students. Correlations between odd and even items ranged from .75-.85 for Form V and .77-.88 for Form VI.

Test-retest reliability was established using Form V of the PEP through two studies: (1) a sample of 19 junior college students with a three month interval

between testings and (2) a sample of 43 first year university students with a nine month interval. Correlations between testings on the three dimensions ranged from .61-.78 on the three month interval study and from .66-.87 on the nine month study.

Inter-correlations among the dimensions were conducted in a study that involved 1,242 junior college students. Correlations among dimensions ranged from .51 to .63. The researchers interpreted the considerable dependence among factors as to be expected since the factors were all epistemic styles but that there was enough independence reflected support that these three dimensions were meaningful.

Validity of the PEP

The basic question regarding concurrent validity in regards to this instrument was, "Does knowledge of a person's score on the PEP allow an accurate estimation of that person's epistemic style?" The authors of this instrument used general experience as their criterion measure and assumed that biologists were most likely to have a commitment to empirical epistemology, novelist and poets to metaphoric epistemology, and mathematicians and theoretical physicists to rational epistemology. Validity of these relationships was supported in the piloting of the first version of the PEP (Royce, Mos, & Kearsley, 1975). Further evidence supported the concurrent validity of the PEP when the PEP (Revised Experimental Form IV) was administered to professional persons engaged in mathematics-philosophy, geophysics, experimental psychology and speech drama. Kearsley repeated these studies in the piloting of the PEP (Revised Form VI) with graduate students in the life science, the humanities, and the analytical sciences. As expected, the humanities students had the highest mean on the metaphorical dimension, the life science students on the empirical dimension, and the analytical science students (mathematics and theoretical

physics) on the rational dimension. It is important to note that these differences had less than a probability of .05 that the difference between dimensions would have occurred by chance. Kearsley also found a significant difference ($p < .001$, t-test) between females and males on the metaphorism scale but no differences on the empirical and rational scales.

Construct validity is determined by examining the degree to which certain constructs account for performance (Payne & McMorris, 1967). A measure of construct validity can be found by correlating the test to tests measuring the same or similar constructs. Mos completed two studies with Form III of the PEP that correlated scores on the PEP, the Allport-Vernon-Linzey Study of Values, the Myers-Briggs Type Indicator, and the Edwards Social Desirability Scale. The first study was done with 86 first year graduate students and the second with a different population of 98 first year students. Correlations were modest and are reflected in a table from the PEP manual (Royce, Mos, & Kearsley, 1975). This is shown on the Table 2 on the following page.

The authors indicated that correlations between the Psycho-Epistemological dimensions and the types on the Myers-Briggs were not well defined. Correlations between the PEP and the Edwards Social Desirability Scale were also small and negative. Actual data from these correlational studies were not represented in the 1975 Manual of the PEP.

In the PEP manual, the validity study that had the most import for this study of the relationship between epistemic style and evaluation practice was conducted by Zelhart and Wargo and reported at the 1971 Rocky Mountain Psychological Association Annual Meeting. The study related the PEP (Form V) and three derived scales to the practice orientation of psychotherapists and the demographic

characteristics and outcomes of clients. The derived scales were obtained from a factor analysis study of the PEP. Correlations were made between the original PEP dimensions and the derived factors with demographic and patient outcome scores. This information is shown below in Table 2.

Table 2
Correlations of Epistemic Style to Other Constructs

	Rationalism	Metaphorism	Empiricism
Theoretical	.26	-.50	.36
Economic	-.04	-.30	.40
Aesthetic	-.23	.48	-.36
Social	-.07	.23	-.23
Political	.09	-.31	-.31
Religious	.01	.31	-.40
Introversion (E-I Scale)	-.01	-.17	-.20
Intuitionism (S-N Scale)	-.01	-.17	-.20
Feeling (T-F Scale)	-.58	.54	-.12
Perceiving (J-P Scale)	-.14	.25	.20
Edwards Social Desirability	-.04	-.37	-.40

Findings suggested an interaction between therapist epistemologies and the demographic characteristics of the client with respect to change due to psychotherapy. The researchers proposed that therapists with metaphorical epistemologies may be

less able to help clients with particular characteristics than therapists with other dominant epistemologies.

Population and Sample Selection for the PEP

In this study the PEP was administered to a random sample of 60 members of the Michigan Association of Evaluators (MAE) from the total population of 107 pre-registered attendees of the 1999 Annual conference of the Michigan Association for Evaluation. This population was chosen on the premise that it represented individuals who were active in the field of evaluation. A random sample chart was used.

Administration of the Instrument

Each evaluator in the sample was sent a packet that included (a) a personal letter delineating the purposes and methods of the study, assuring confidentiality, and requesting his/her participation (See Appendix D); (b) the PEP, along with questions asking for demographic information; (c) an informed consent form for the administration of the PEP; and (d) a self-addressed stamped envelope to facilitate a high rate of response. Each instrument was coded to enable follow-up of non-respondents. Ten days after the sending of the instrument, a follow-up reminder was sent to non-respondents. Ten days after the reminder, a self-addressed post-card was sent requesting information regarding the reasons they chose not to participate.

The instrument was self-administered; complete instructions were given on the cover page of the test booklet. The instrument could be completed in approximately 15-20 minutes. Participants were encouraged to answer all questions.

If they desired to qualify any of their answers, space was provided at the end of the test to do so.

The PEP produces separate scores on three dimensions representing the three epistemic styles: (1) empirical, (2) rational, and (3) metaphoric. The value of items in one dimension does not affect the values of items in another dimension, making each of these scales independent of the other scales. There are 90 items on the instrument, 30 measuring each of the dimensions. Participants indicated their agreement to each item on a five-point scale (complete disagreement-1, moderate disagreement- 2, neutral-3, moderate agreement-4, and complete agreement-5). Instruments were hand-scored by the researcher

Analysis of Instrument

Reliability

The main purpose for the use of the PEP in this study is to ascertain the epistemic style of the evaluators in order that the relationship of epistemic style to evaluation practice can be examined. This instrument was normed and used primarily with college students (including junior college, first-year university students, and graduate students in Alberta, Canada). I make the assumption that this population has important differences in personal characteristics from my particular study population. For this reason, a large sample was chosen in order to examine the reliability with this particular population of professional evaluators. This was accomplished by measuring the internal-consistency using Cronbach's coefficient alpha. (Ysseldyke, 1981). The coefficient alpha is based on the average split-half correlation based on all possible divisions of a test into two parts. All data were entered into a database

using SPSS Software (SPSS, 1998). Analysis was conducted with SPSS.

Reliability of the PEP With Professional Evaluators

Coefficient alpha's were calculated for each of the scales—metaphoric, rational, and empirical. This information is found in Table 3 below.

Table 3
Reliability of Psycho-Epistemological Profile

	Metaphoric	Rational	Empirical	Total Test
Coefficient Alpha	.77	.76	.79	.88
Number of Items	30	30	30	90
Mean (N= 38)	3.64	3.43	3.24	3.43
Standard deviation	.62	.72	.60	.66

The alpha coefficients, ranging from .76 to .79 are reasonably high. There were 39 subjects who completed the instrument to be used in this study, a response rate of 65% from the number in the original sample. The 39 subjects that were used in the reliability study, represent 36% of the population selected for study. Of those that did not respond, three indicated they were no longer practicing evaluation, and nine indicated they had lack of time to participate. This information about non-respondents may mean there was a group (or type) of evaluators that was not included

in the study. This may have been those evaluators having a dominant empirical style.

Subjects chosen for interviews had both commonalities and differences. All presently work as professional evaluators or evaluation consultants and are women. The difference and commonalities within each style will be noted in the second session of the analysis that extrapolates an evaluator typology based on epistemic style.

Hierarchy of Epistemic Style

For each subject, scores were totaled for each of the three scales (empirical, rational, and metaphoric). This yielded a hierarchy of epistemic styles with the highest score being the dominant epistemic style of the individual. The next section describes how subjects were chosen from each of the dominant epistemic styles to participate in the second portion of the study that relates these styles to evaluation practice.

Relationship of Epistemic Style to Evaluation Practice

Description of Methods

I chose qualitative research methods to study evaluation practice in order to elicit rich data that not only yielded the type of activities evaluators use in their practice, but also to uncover assumptions underlying their practice. My goal was to elicit participants' perceptions and understandings of their evaluation practice and examine whether these descriptions of practice were similar to evaluators with the same dominant epistemic style.

The primary method of gathering data on evaluation practice was by interviews of approximately forty minutes in length that consisted of standardized

open-ended questions and an interview- guided process (Patton, 1987). The interview model I chose for use in this study was the “Developmental Research Sequence (D.R.S.) Method” developed by James Spradley (1979). It is a systematic method of doing ethnography characterized by a sequence of tasks. As this study was not a comprehensive ethnography, it utilized a portion of the sequence related to interviewing and data-analysis.

The majority of the interview consisted of a “guided-interview” process in which topics were determined in advance but without the precise wording of the standardized questions. Topics to be addressed included methods of developing evaluation plans, methods of data gathering, involvement of stakeholders/clients, methods of communication and/or reporting, involvement in the utilization of evaluation data, and ethical dilemmas.

Ethnographic questions were asked that included descriptive and structural questions (Spradley, 1979). Descriptive questions were designed to gather a general picture of an individual’s evaluation practice. An example was, “Could you describe a typical day in your work?” Structural questions were those questions that allowed me to understand how the respondent organized their knowledge. “What are the stages you go through in developing an evaluation plan?” was an example.

Development of the Interview Protocol

The interview protocol (See Appendix F) was developed by reviewing the current literature in the field of evaluation, reflection on personal evaluation experience, and feedback from professional evaluators. Feedback was gathered from three evaluators working in Southwestern Michigan. Evaluators were asked to comment on the relevancy of questions to the topic of study, appropriateness of

language, and comprehensiveness of the protocol

Selection of Sample for Interviews

A purposive sample was chosen from the population of evaluators who completed the PEP. Nine respondents were chosen. It was the intent to pick three evaluators from each of the empirical, rational, and metaphoric epistemic styles.

At the conclusion of the reliability study of the Psycho-Epistemological Profile, it turned out there was only one evaluator that represented the empirical style. The evaluator representing the empirical style was not available for interviewing in this study. In examining the hierarchy of epistemic styles, it became apparent that there were a notable number who had less than a five point spread between the three styles. I was curious to see if this might represent an entirely separate style that was significantly different from either rational or metaphoric styles. Three subjects were selected from this group of seven. In this study, this group was identified as Eclectic. This category reflects those individuals who did not exhibit a dominant epistemic style.

The sample was selected from the highest scores in each dominant style, with the rationale being that these individuals would represent strong examples of the style. Interviewees represented a range of experience (less than 2 to over 10 years) and came from a variety of types of organizations: universities, a community college, non-profit foundations, a non-profit service organization, and a public service agency. All interviewees were women.

Selected evaluators were first contacted by mail. This letter described the purpose of the interview and an overview of the interview topics. A week later, these evaluators were contacted by telephone, requesting their participation. A description

of the characteristics of the sample is shown in Table 4.

Table 4
Description of the Interview Sample

Respondent	Years of Experience	Type of Organization	Empirical Score	Rational Score	Metaphoric Score
Metaphoric A	Less than 5	Non-Profit Foundation	100	107	132
Metaphoric B	5-10	College/ University	91	96	115
Metaphoric C	5-10	College/ University	103	108	126
Rational A	5-10	Non-Profit Organization	91	111	101
Rational B	Less than 5	Non-Profit Organization	111	114	101
Rational C	5-10	College/ University	95	106	99
Eclectic A	Over 10	Non-profit Organization	98	100	103
Eclectic B	Less than 5	College/ University	108	109	106
Eclectic C	5-10	Non-Profit	117	117	118

Interview Process

Interviews were conducted using a portable tape-recorder in closed settings and were approximately forty minutes in length with only the interviewer and the

interviewee present. Notes on interview content were taken to provide back-up in case of mechanical failure of the tape-recorder. Two scheduled interviews needed to be changed to telephone interviews because of a personal emergency of one interviewee. Both interviews were scheduled initially on the same day, 200 miles from my home. When one interviewee requested a phone interview to take the place of a personal interview, the other interviewee agreed to the same process. For each phone interview, the interviewee was aware I was using a phone recorder. Each tape was labeled with the date of the interview and an identification number (ID#) which was a two digit number. Prior to the interview, each respondent was given an overview of the topics to be discussed and a description of measures taken to assure confidentiality. Interviewees were referred to only by ID# and the ID# will be known only to the interviewer and the respondent. All identifying information on tape (i.e. place names, individuals' names) were disguised in both the transcript and the final research paper and a copy of the transcript was available to the interviewees on their request. After each interview, tapes were transcribed by myself and did not leave my possession. Within four weeks, each respondent was provided with a copy of the analysis of her epistemic style. Interviewees were given the opportunity to respond or object to the analysis. Analysis was confirmed by interviewees.

Interview Data-Analysis Overview

Spradley describes the first step in the analysis as componential analysis (1979) which means clarification of components of meaning among symbols or phrases. It includes the study of semantic relationships. For example, a discussion about utilization-focused evaluation, the respondent may identify the semantic relationship by a comment such as "evaluations must be utilization-focused or there is

no justification for conducting the evaluation.” Or a respondent can explain his/her components of meaning in the term “evaluator”. This is illustrated in Figure 1.

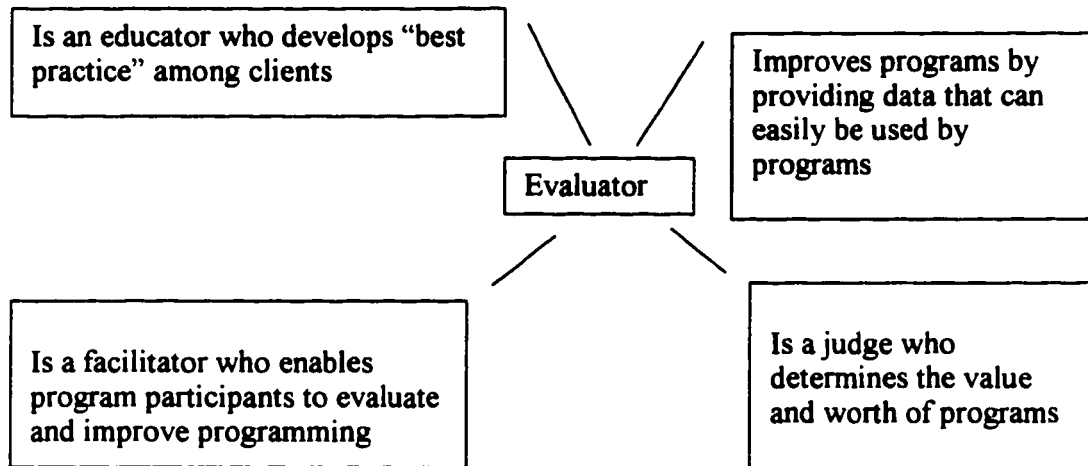


Figure 1. Components of Meaning.

The next step in the analysis of data was to identify themes and patterns within each of the domains identified in the interview protocol: (a) purpose of program evaluation, (b) role of the evaluator, (c) development of evaluation plans, (d) methods of data gathering, (e) involvement of stakeholders/clients in the evaluation practice, (f) methods of communication and reporting, (g) involvement in utilization of evaluation results, and (h) ethical dilemmas. This process of identifying patterns was done solely by the researcher without the aid of computer software. When patterns were identified for each of the epistemic styles, a member-check (Lincoln & Guba, 1985) was conducted by providing the analysis for each of the respondents. They were asked to comment regarding their agreement or disagreement with the analysis. This was done to establish credibility, which means the reconstructions of reality were deemed credible by the interviewees.

Establishment of Reliability and Validity of Interviews

Egon Guba (1981) described credibility and reliability in naturalistic inquiry as credibility, transferability, dependability, and confirmability of the study. As described previously, credibility was sought by conducting “member checks.” In presenting the findings to informants, no personal or controversial information that allowed evaluators to identify each other was presented. The focus was on whether the informants found the analysis to be consistent with the world of evaluation practice.

Transferability of a study was established by providing descriptions of the context of the study. These descriptions allow the reader to make judgments as to whether the analysis fits other particular contexts. Dependability and confirmability of this study was ensured by the detailed records of notes, interview protocols, data analysis process, and member checking.

Summary

The research questions for this study were:

- 1. What is the relationship between evaluators’ personal epistemic style and their evaluation practice?**
- 2. Does the evaluator’s epistemic style provide a foundation for evaluation practice?**

This study grew out of my own autobiography as an educator and evaluator (See Appendix A). It is impossible for anyone to see from exactly my perspective as each reader has his/her own story from which he/she frames the world. Therefore this study illuminates as much as possible my own bias, as well as process, in order that the findings are comprehensible and of value to others.

CHAPTER IV

DATA ANALYSIS

Personal Style

This research study was based on the premise that personal theories influence one's evaluation practice. It is appropriate therefore, that I explicate my own epistemic style and examine how my hierarchy of epistemic style, combined with other experiences, have affected my own analysis of the collected data.

At the very beginning of the study, I completed the Psycho-Epistemological Profile that was to be completed by the subjects of this study. My dominant epistemic style is metaphoric. My secondary style is rational. There was a 46 point spread from the score of my least dominant style (empirical) to my most dominant style (metaphoric). This is in contrast to the average point spread of 15 points among epistemic styles across all participants. This finding seems congruent with how I see, learn, and interpret the world. I rely highly on my intuitive senses to make sense of the world. I am most comfortable in expressing myself creatively with art, poetry, and stories. When I worked as a teacher, this personal characteristic was strongly evident in my classroom and in my interactions with students. The walls, and sometimes the ceilings, were always filled with art, we created our own plays for holidays, and generally had a "story time" every day when I was teaching at the elementary school level. In my role as a program evaluator, my rational style is most evident. On entering my office, you will see organizational charts, a database on the computer screen, and files filled with evaluation criteria. Data are collected and

reported in a very systematic way. This is actually less oppositional than it may seem. The systematic collection of data is the modus operandi of the organization for which I work. However, it is not entirely antagonistic to my style. I continually look for opportunities in which I can insert holistic approaches to data gathering and reporting such as the use of case studies or as I did in one case, a visual montage in the form of a large poster. My intuitive sense is evident in the way I deal with clients to make the evaluation experience a collaborative event rather than being an experience where we, as external evaluators, come to impose our criteria and judgments. I strive to include the views of our clients and the relevant stakeholders in the evaluation.

My work as a program evaluator is combined with my graduate education experiences at Western Michigan University. My curriculum in Educational Evaluation, Measurement, and Research was heavily grounded in systematic inquiry using quantitative methods but did include classes in naturalistic inquiry, policy analysis, and administrative analysis that were grounded in qualitative research. From this experience, I have learned the value of combining qualitative and quantitative approaches for a more holistic, comprehensive picture of meaning. In this study, I have used such an approach. Salomen (1991) explained this as a systemic approach to understanding the interaction of variables in a complex environment. The first part of my study necessitated the identification of subjects representing different epistemic styles. One approach to this could have been to use intensive personal interviews of a number of subjects, inquiring how subjects established their own criteria for knowledge, and discovering the characteristics of their personal theories of knowledge. I rejected this approach because I am not sufficiently grounded in psychology, epistemology, and clinical interviewing. As an

alternative, I selected an instrument that was developed by scholars with such expertise. This instrument was thoroughly tested for reliability and validity by the original developers of the instrument. Because their studies were conducted with graduate and undergraduate students, I felt the need to extend the reliability test to my chosen subjects—professional evaluators. From this reliability study, which is described in Chapter III, I was also able to choose subjects for the second part of my study examining evaluation practices.

I have approached analysis in a way that reflects my preferred epistemic styles metaphoric and rational. The next section attempts to answer the first research question, “Do evaluators representing different epistemic styles exhibit differences in evaluation practice?”, by systematically looking at the differences in pre-defined categories of responses across subjects representing three epistemic styles. These categories were taken from the main categories of discussion in the interview protocol. This is a more rational approach that is working from experience (the interview protocol) and defined objectives.

The third section of this chapter focuses on answering the second research question, “Does the evaluator’s epistemic style provide a foundation for evaluation practice?” By searching for major patterns, trends, and themes within each style, a profile was developed of each of the styles. Through repeated reading of the interview transcripts, themes were identified that were similar across the three interviewees within a style. When a theme was qualitatively different in one respondent, further analysis of word choices was done to try to extrapolate where the differences (if any) resided. These themes were then examined in each of the other two styles to determine whether there was evidence of a similar pattern. This is a more metaphoric approach that is based in “making sense” of the data. In a typical

metaphoric way, this section is not entirely committed to this one approach but attempts to look at the data from different angles to bring meaning.

Relationship Between Epistemic Style and Evaluation Practice

Research Question #1

The first research question of this was, “Do evaluators representing different epistemic styles exhibit differences in evaluation practice?” Analysis of interviews provided evidence that there were both differences and similarities among evaluators representing different epistemic styles.

Analysis of Interviews

After interviews were transcribed, a table was created that summarized responses across pre-determined categories linked to the interview protocol (See, Table 9, Appendix H). These categories were (a) typical day, (b) theoretical influences, (c) preferred entrance, (d) development of evaluation plan, (e) audiences for evaluation, (f) role of evaluator, (g) role of clients, (h) role of stakeholders, (i) primary methods of data collection, (j) planning data analysis, (k) types of analyses, (l) communication of reports, (m) sharing of draft reports, (n) recommendations, (o) use of data, (p) ethical dilemmas, and (q) case study analysis. Each column represents an epistemic style: metaphoric, eclectic, or rational. By examining summaries across columns, differences and similarities between styles were found. Each of these categories are described along with the findings at the end of the discussion regarding differences among styles. In describing similarities and differences, evaluators will be identified by their epistemic style and a letter. For example, the metaphorics will be called Metaphoric A, Metaphoric B, and

Metaphoric C. When a subject is directly quoted she will be identified in this way (Rational B, p.5). The “p.5” refers to page five of Rational B’s transcript from the interview.

Differences Among Epistemic Styles

There were differences among epistemic styles in seven categories of practice. These were (1) typical day, (2) theoretical influences, (3) role of the evaluator, (4) role of clients, (5) role of stakeholders, (6) planning data analysis, and (7) types of analyses. Each of these differences is described in this section.

Typical Day

To gain an understanding of how evaluators organized their day, I asked evaluators to describe a typical day in their work. Metaphorics had difficulty describing their days. “A typical day is atypical,” explained Metaphoric A. This group seemed to thrive on variety and ambiguous definitions of the job.

There are a couple of streams of things and a typical day is atypical. (Metaphoric A, p. 1)

It might be one project, it might be more. It might be juggling communications from multiple projects. (Metaphoric B, p. 1)

I’m not only doing evaluation.....they ask for my input on not only evaluation but also for program development and program change and all those kind of things. (Metaphoric C, p. 1)

Rationals were task oriented and described activities such as collecting data, managing data-bases, generating reports, selecting instruments, determining methodologies, etc.

My job is to collect data. (Rational A, p. 1)

I do all the grant process. (Rational B, p. 1)

The Eclectics leaned more to the Rationals' descriptions but also described a process that sounded like exploring.

I will spend an hour on the Internet looking for grants that are open, checking the news from the Ed Dept for any new information.
(Eclectic A, p. 1)

....trying to understand neighborhood perspectives of issues as well as secondary data. (Eclectic B, p. 1)

So I'm looking for relative information and trying to put it together.
(Eclectic C, p. 1)

In summary, the Rationals were task-oriented, the Metaphorics thrived in ambiguous and varied, frameworks, and the Eclectics were both task-oriented and explorers.

Theoretical Influences

Evaluators were asked if they could name theoretical influences on their practice. They were told they could name specific theories, individuals, or experiences. The purpose was to gather evidence about how they used theory in their practice. Theoretical influences had some similarities across styles but there was evidence of significant differences. Similarities were primarily in the use of Michael Quinn Patton's Utilization-Focused Theory (1997) and an outcomes-based evaluation model used by the United Way. Both the Metaphorics and the Eclectics cited these.

Rationals seemed to operate from experience or well-established theory. Theories that influenced their evaluation practices came from experience or were substantiated by reliable sources. Rational B spoke of using psychological theories such as those of Piaget and Maslow to facilitate research and assessment. In addition she looked for theories to support the perspectives of participants. It seems that this evaluator is using theories to better understand the people with whom she is working

rather than a base to develop actual practices. This also reflected her commitment to effectively include stakeholders in the evaluation process. Rational A used her experiences as a teacher.

The biggest influence in my life is in my years as a teacher. There the influence really came from the students. I came to understand not only my field but also how people approach knowledge and how to convey the knowledge to them. Not one person...but that dynamic. (Rational A, p. 2)

Rational B could not identify any influences. This may be a reflection of this evaluator being new to the field of evaluation and not having formal evaluation training. She did mention at another point in the interview, how useful professional workshops were to her.

Whenever there are workshops, I try to attend whatever I can in the evaluation arena, but it is still a new area for me and I'm still learning some of the dialogue. (Rational B, p. 1)

Two of the Metaphorics relied on similar evaluation theories. Both identified with Michael Quinn Patton's Utilization-Focused theories (1997) and David Fetterman's Empowerment Theory (1996). These theories both place heavy emphasis on involvement of stakeholders in the evaluation. Metaphoric B cited Robert Stake as an important influence. Stake's model of congruence looked at whether the activities in a program (the process used) match with the intent of the program. This responsive model of evaluation focused on responding to the participants' actions in carrying out the program. Metaphoric C referred to United Way's outcomes-based model. Both of the models seem to look for a fit among theory, process, and outcome. The third metaphoric evaluator did not name specific theories but referred to this same outcomes-based model that links good questions with process and outcomes. She referred to this as "evaluation for learning" and identified it as the only "out-loud" framework that she has. Throughout her interview though she

alluded to using her intuitive sense to make meaning.

I think I listen and file away and pull back-out and sort.....I think I'm good at sorting together towards ahas! (Metaphoric A, p. 13)

The Eclectics did not seem to have a need for verified, defined theory as did Rationals nor did they rely strictly on experience. They tended to combine experience with theory. Eclectic A, named theorists whom she had met who deeply influenced her.

Patton. When I heard him speak, the things he said ring true to me and I try to do that bringing of clients into the evaluation so that it's a development process rather than me doing it for them. (Eclectic A, p. 1)

Eclectic C referred to Egon Guba and her experience using his model in a research project in one of her graduate classes. Eclectic B referred to models commonly used in her type of agency: Wholey's Logic Model, Patton's Utilization-Focused theory and organizational models through the independent sector.

In summary, the Rationals operated from experience or well-established theory; the Metaphorics attempted to link theory, process, and outcomes; and the Eclectics tended to combine experience with theory. There is more to be learned from an examination of the types of theories chosen by representatives of the three epistemic styles. This will be explicated further in the last section of this chapter.

Role of the Evaluator

Evaluators were asked to talk about their ideal role as an evaluator. Each evaluator was given a stack of cards identifying various evaluator roles. First they were asked to discard those roles that did not describe their ideal role. Then they were asked to rank order the cards they chose. They were also encouraged to write down any additional roles they felt were important to them. I then asked subjects to

explain their choices. The cards read: facilitator, expert, investigator, patron, translator, judge, mediator, critic, consultant, connoisseur, power merchant, interpreter, educator, describer, negotiator, and change agent. These roles were compiled from the literature (Affholter, 1994; Eisner, 1991; Guba & Lincoln, 1989; Patton, 1980; Patton, 1997; and Worthen & Sanders, 1987). The purpose of using the cards was to initiate conversation about ideal evaluator roles.

The most significant differences and similarities occurred in the first three cards chosen as primary roles. These choices are shown in Table 5.

Table 5
Primary Roles of Evaluators

	Metaphoric	Rational	Eclectic
Respondent A	1) change agent 2) facilitator coach 3) interpreter	1) educator 2) investigator interpreter 3) facilitator change agent	1) facilitator 2) coach 3) educator change agent
Respondent B	1) facilitator 2) coach 3) change agent	1) educator 2) investigator 3) facilitator coach	1) facilitator 2) investigator 3) translator coach change agent
Respondent C	1) leader change agent 2) facilitator/coach educator/ collaborator	1) educator 2) facilitator 3) coach	1) change agent 2) facilitator 3) investigator

Among these three cards, all three groups chose facilitator and coach as very important roles. The difference is in how they interpret the words. In describing

coach and facilitator Eclectic A said:

I like these words. Because they relate to other things other than evaluation. That notion of there being a team, teamwork to get where you want to go. (Eclectic A, p. 5)

Rationals seemed to interpret these same words as a way to get the work of evaluation done. Rational B in talking about coaching said:

It's a role I play with staff. I see myself as designing the evaluation tool but the staff who are actually administering the different programs... coaching them, that they actually administer the evaluation tool and they buy in to it and they're part of the process. (Rational B, p. 4)

The Metaphorics seemed to interpret coach and facilitator as helping roles.

Facilitator and coach are the roles I like to play in an evaluation. I like to help people do, as well and thoughtfully as possible, what it is that they are trying to do. (Metaphoric B, p 10)

It doesn't seem to me that I have a role that imposes things on people but rather brings all the possibilities to the table and changes people and their choices. (Metaphoric A, p 9)

Changing people is a big priority for Metaphorics and Eclectics and constitutes a major difference between the primary choices of Metaphorics, Eclectics, and Rationals. Metaphorics and Eclectics chose change agent among their first three primary roles.

And it's a change agent toward improvement. I'm a big believer that we can learn things and we can improve and that's a real opportunity through evaluation. (Metaphoric A, p. 9)

Change agent. That will be my ultimate goal if needed. Facilitator, that goes along with change agent. Make it possible to happen. (Eclectic

Rationals chose educators as their most primary role. This can be considered to be quite different from change agent, depending on how it is interpreted.

Educator. That goes back to my background. The primary role of this job is statistics but what statistics do is educate. Where are we, what are we going to do with it? So that's my primary role of all those things. (Rational A, p. 5)

Educator. I think its educating the players involved...staff, board. Educating them as to the benefits and not only of the importance of evaluation, but the benefits to the agency for future planning, strategic planning, but also educating the staff on how evaluation can help you do your job. (Rational B, p. 6)

Rational C linked educator to facilitator, coach, and change agent. Her goal is to educate people about their own research and evaluation.

To give them power to think about their research and answer their own research questions. So that's why the facilitator and the coach come out. Working together with other people and letting them, you know, research. (Rational C, p. 8)

After the primary roles, the only noticeable pattern was one of similarity in what roles were discarded. All evaluators discarded the following roles: expert, judge, connoisseur, and power merchant. This is an interesting finding when one considers the traditional definition of an evaluator as being one who determines the significance, worth, or condition by careful appraisal and study (Merriam Webster's Collegiate Dictionary, 1997). Choices reflected a more transformative type of evaluator. This was described more specifically by Metaphoric C when she added the leader role to her preferences.

I don't believe that we are the only leader but that we fulfill a leadership role when we do evaluation. In giving guidance. It's more of a transformational leadership style when we do evaluation. I see myself doing that in the program management part. Or asking questions, raising questions, in a sense empowering people to take charge and change things. (Metaphoric C, p.10)

In summary, all three of the Epistemic styles highly valued the facilitator and coaching roles. However, acting as a change agent was more important to the Metaphorics and Eclectics. Acting as an educator was more important to the

Rationals. How each epistemic styles defined their roles will be further illuminated in the last section of this chapter.

The Roles of Clients

Informants were asked about what roles they prefer clients and stakeholders take in a program evaluation. In regards to clients, metaphorics were consistent in their responses.

Actively, early, often, thoughtfully. (Metaphoric B, p. 11)

Clients, if you will, need to take the lead role. (Metaphoric A, p. 11)

I would like them to be evaluators, too. (Metaphoric C, p. 11)

They did express though that doing this is not as easy as it sounds.

So pulling people into evaluation is sometimes just as hard as remembering that they ought to have been there to begin with. (Metaphoric A, p. 11)

Eclectics varied considerably in their responses. Eclectic C who worked for a foundation in a role as a consultant to agencies planning evaluations, encourages them to include community engagement. Eclectic A and B expressed opposite viewpoints.

A learning role. An active learning role so they're not doing this because they have to. They're not doing it with the intent of put it on a shelf. Or are frightened by it but they really want to make the most of what an evaluation can do. (Eclectic A, p. 6)

Truth be told, I would kind of like them to inform me of what they are looking for or what they are seeing as important and facilitate getting the information to me that I need, and then step back and let me do my work. (Eclectic C, p. 6)

Rationals spoke more of what actually happens rather than what they would hope to happen. Consequently their answers may be more closely linked to the organizational culture in which they work.

I don't control any of that. (Rational A, p. 9)

Paid staff. They have taken a minimal role in the process. I think they see the value of evaluation, the benefits of evaluation and why its necessary for many funders, but I think the perception is that I go and do it. (Rational B, p. 9)

Include clients. Especially in the A Program. They were involved in each step of the process. It was really participant centered. They became participants of the whole research process. (Rational C, p. 3)

In summary, Metaphorics were most likely to involve clients in active roles, while the Eclectics and Rationals varied in attitudes about stakeholder involvement.

The Role of Stakeholders

The preferred role that subjects would like stakeholders to take in evaluations follows a similar pattern as to what they prefer for clients. Metaphorics were in agreement that stakeholders should take an active part in the evaluation even though this is sometimes difficult.

The thing I continue to be challenged by is how do you include and involve stakeholders, clients, etc. when all of the societal pressures say get it done, get it done fast, make it look good, what do those people know anyway, move on the next thing. I have a lot of commitment to that and yet I still struggle with how do you make that real. (Metaphoric A, p. 12)

The Eclectics varied from taking an active role, to having a limited role.

It's hard.....Sometimes I find those people are hard to deal with because they are so...they're on a roll of making things happen rather than thinking of a broader picture of education. I think that a lot of times they let other things interfere with being able to think about the system rather than about their child or about the situation. (Eclectic A, p. 6)

I would like them to have an active role. I see them as the source of information. So I would like them to have an active role in shaping

what I should be looking for. And possibly helping me interpret results. (Eclectic B, p. 6)

Rationals agreed that it was very difficult to involve stakeholders in an evaluation. One Rational reported that in her job she had little contact with stakeholders and another Rational reported she would like to involve them more.

In summary, Metaphorics were most likely to involve stakeholders in active roles, the Eclectics varied in attitudes about stakeholder involvement, and the Rationals would like to involve stakeholders more than they were presently.

Planning Data Analysis

Informants were asked to describe at what point they determined their data analysis. Metaphorics and Eclectics agreed that this was an evolving process that was determined throughout a formative evaluation process. This reflected the type of evaluations they liked to do. Eclectic A specified that as data collection methods were determined, she decided on her analyses. The Rationals were not in agreement but did set definite times for determining analyses. One said “before,” one said “after,” and one said “before and after.”

I try to plan my analysis before. You know when you are asking research questions in the evaluation you need to decide what level of measurement you need, although measuring will also impact your data analysis and theoretically you shouldn't data mine. I don't try to data mine. (Rational C, p. 11)

Types of Analyses

Evaluators were asked to identify what types of analyses they usually do. Subjects were given a set of cards on which various analyses were written. The choices were (a) descriptive—quantitative methods such as sums, percentages, averages; (b) descriptive—content analysis of qualitative data such as interviews,

focus groups, and open-ended questions on surveys; (c) descriptive—case studies; (d) study of relationships using cross tabulations or correlations; and (e) inferential statistics. Subjects were also encouraged to add cards. The subjects were requested to then rank the cards according to preference and talk about their choices.

The most significant difference was that all of the Metaphorics ranked case studies as the most preferred choice.

I think that I would say that extrapolating case studies and the good story is something that is really important to me. (Metaphoric A, p. 15)

The two I prefer are descriptive case studies and content analysis. Because I think that is rooted in not only my perception of something that's useful but its rooted in the perceptions of interviewees as near as I can determine what they are. (Metaphoric B, p. 15)

The metaphorics also all chose content analysis of qualitative data as their second choice. The Rationals and the Eclectics did not show a pattern of agreement in their choices other than use of case studies (if it were chosen at all) was their least preferred method of analysis.

In summary, the Metaphorics preferred use of case studies, while this was the least favored among Eclectics and Rationals.

Summary of Differences Among Epistemic Styles

There were differences among epistemic styles according to how evaluators described their typical day, their theoretical influences, the evaluator roles they favored, the roles they preferred for clients and stakeholders, when they planned their data analyses and the type of analyses they chose to do.

Categories of No Difference

Notable differences were not found between styles in eight categories. These categories were: (1) time of entrance into an evaluation, (2) development of evaluation plans, (3) purposes of evaluation, (4) communication of reports, (5) sharing of draft reports, (6) including recommendations in reports, (7) use of data, and (8) ethical dilemmas. The categories in which no pattern of differences could be found between different epistemic styles, can be classified into two groups: (1) those categories that evaluators responded in very similar ways across all epistemic styles and (2) those categories that showed no pattern of differences between epistemic styles and showed evidence of dissimilarity within an epistemic style.

Similarities Across Styles

The five categories in which evaluators responded in similar ways across epistemic styles included: (1) time of entrance into an evaluation, (2) purposes of evaluation, (3) sharing of draft reports, (4) including recommendations in reports, and (5) use of data.

Time of Entrance Into an Evaluation. All evaluators preferred to enter an evaluation process at the beginning of a program. Although this was the ideal, some indicated circumstances when it was appropriate to enter at other points. Eclectic C felt comfortable entering at any point.

Sharing of Draft Reports. With the exception of Eclectic C, evaluators believed in sharing draft reports with clients to solicit feedback before final reports were presented. The one exception made an interesting comment that reflected her concern that sharing could possibly compromise the integrity of the report.

I prefer not. I have a problem with having to accommodate to, I mean having to alter something to please the program managers or stakeholders. (Eclectic C, p. 9)

Including Recommendations in Reports and Use of Data. Helping clients use data and making recommendations for improvement were linked for most evaluators. One of the Eclectics qualified her response regarding recommendations.

“In what I’ve come to now...is having those recommendations come from them...so they are the ones thinking about solutions. I guess recommendations and solutions might be different. I mean I could recommend that they study this or that they more fully understand whatever issue comes out of there but as far as solving problems or changing a program, I think those are things they have to do and I can certainly facilitate that. (Eclectic A, p. 9)

Purposes of Evaluation. A card sorting exercise was used to determine what evaluators thought should be the primary purposes of program evaluation. These choices were based on the purposes of evaluation found in the literature (Weiss, 1998; Worthen & Sanders, 1987; and Guba & Lincoln, 1981). Choices on the cards were (a) to document whether program objectives have been met, (b) to enable program staff to make changes that improve program effectiveness, (c) to foster accountability, (d) to determine the effectiveness of programs for participants, (e) to determine whether implementation and outcomes of the program “fit” with the theories the program is based on, and (f) to facilitate program management. Evaluators were asked to rank order the purposes they felt were important, add any additional purposes they cared to, and to then discuss their choices. For all epistemic styles, primary choices were to enable program staff to make changes that improve program effectiveness and to determine the effectiveness of programs for participants. In addition, two Metaphoric evaluators chose foster accountability as the most primary purpose.

It's one of those things I'm really struggling with at the moment so it is taking a lot of my thinking time in terms of identifying ways we can hold ourselves accountable primarily to the service population or external audiences. (Metaphoric A, p. 3)

One of the things that I have come across lately is really something that is of concern. When you work with a program and you find out there is no program. (Metaphoric C, p. 4)

Another type of accountability issue came up for Rational A. Her concern was regarding the need for conducting needs assessment (a purpose for evaluation) prior to developing a program.

The first thing I want to know is do you need a program? I mean do you have a smoking problem? If you don't have a smoking problem, why do you need a program? If you need a program, is what your program going to do have an impact on this? (Rational A, p. 3)

Primarily, in this study, evaluators were concerned that their evaluations are used for improving programs and peoples' lives.

Let's do something that's worthwhile, that's the main priority I think. (Rational C, p. 4)

Summary of Similarities

In summary, evaluators across epistemic styles generally agreed that (a) they preferred to begin evaluation at the initiation of a program, (b) sharing drafts with clients to elicit feedback was valuable, (c) including recommendations with reports and helping clients use data was important, and (d) enabling staff to make changes that improve program effectiveness and determining the effectiveness of programs for participants were of primary importance.

No Pattern of Difference Between Epistemic Styles

There were four categories that did not show a pattern of difference between

epistemic styles: (1) development of evaluation plans, (2) communication of reports, (3) ethical dilemmas, and (4) analysis of case studies.

Development of Evaluation Plans. Evaluators were asked to describe the typical stages they go through in the development of an evaluation plan. When examining responses to see if there was a pattern of difference by epistemic style, there seemed to be none. But when I considered type of organization or the degree of experience, there were some patterns. Those evaluators working within a university or college identified collaboration with clients as an important step.

I would generally go out and meet with the clients to find out what the questions are. What do they need, what are their interests, what do they want to figure out? (Rational C, p. 3)

Those evaluators with the least program evaluation experience (less than two years), chose focus groups as their initial step.

Communication of Reports. All of the evaluators expressed they did both written reports and presentations. Although some reported they seldom did presentations, this was more a reflection of their organization's culture than it was their personal preference. Two evaluators (a Metaphoric and a Rational) stressed the importance of dialogue and formative reporting throughout the evaluation process. Metaphoric B expressed a quite different preference for communicating findings if she had complete freedom of choice:

In a meeting with stories, or case studies, or videos, for example of real or hypothetical classroom scenes, of real or hypothetical interviews, maybe with an actor stringing together different quotations that are pervasive, or maybe more than that so you could do a counterpoint. Letting them see in less threatening terms what people are saying. (Metaphoric B, p. 16)

Ethical Dilemmas. All but one evaluator felt there were ethical dilemmas in

the field of evaluation. There did not seem to be any evident pattern of the responses according to epistemic style. These dilemmas were: (a) organizations bowing to the wishes of the funder, (b) evaluations not reporting weaknesses as well as strengths, (c) being called in to do an evaluation of a program that doesn't exist in any meaningful way, (d) political aspect of evaluations, (e) evaluation for the sake of accountability without caring about meaning, (f) confidentiality, (g) guarding of information and defense of territory by clients, and (h) conflict of interests (evaluators often paid by organization being evaluated).

Analysis of Case Studies. At the end of the interviews, the evaluators were presented with a short case study (Appendix G) describing a project in early childhood education. Evaluators were asked to describe briefly, how they would approach an evaluation plan of such a project. It was difficult to determine whether there was a pattern of differences between epistemic styles due to the lack of response from some evaluators. Three evaluators did not respond to the case study. One evaluator (a Rational) declined to respond based on the fact that she had a lack of experience to evaluate the type of situation that was presented. Two evaluators (one Rational and one Eclectic) responded that it was not the type of contract they would ever consider, consequently preferred not to respond.

I would be running away from it. Good heavens. I wouldn't take the contract. (Eclectic C, p.10)

The other evaluators varied in whether they would do most of their data-collecting and conferring by phone, by email, by conferences, or on site. There was not a pattern of differences based on epistemic style. Eclectic A and Metaphoric C began with the process of developing logic models to connect theories, practice, and particularly outcomes.

...having the people from the six universities come together for a conference, as you get started, to talk about logic models, program outcomes, and different activities. (Eclectic A, p. 10)

I will follow a product impact approach (Scriven) plus a client outcome based approach using a logic model. (Metaphoric C, p.21)

Rational A began with conducting research on the communities and then visiting sites to get a feel for the project. She was the only evaluator who talked about actually observing the materials being implemented. Her final focus was to evaluate the outcomes of using the instructional materials in six months and then a year. Metaphoric B was the only evaluator that formulated specific evaluation questions. She emphasized the formative nature of her plan.

I would attempt to make interim reports to sites to feed back information and perceptions gathered, both as a reliability check and as formative feedback to the process. (Metaphoric B, p.20)

Metaphoric A stressed that she would not like to approach this as an external evaluator. She preferred to play more of a facilitator's role. She also emphasized that this was an evaluation contract that she would be running away from because of lack of expertise in the educational context of this study.

Summary of Differences and Similarities in Evaluation Practice

A summary of differences and similarities is illustrated in Table 6 on the following page. Descriptions of practice are illustrated in Appendix H. Table 6 illustrates that the major differences are not so much in how evaluators determine purpose or use of evaluation but more in the processes involved in evaluation. Examples of processes are: (a) roles that evaluators, clients, and stakeholders play; and (b) how analysis of data is implemented. There were also differences in theoretical influence among epistemic styles. If the hypothesis that theoretical orientations influence evaluation practice is true, it is not surprising that there are

differences in theoretical influence among epistemic styles.

Table 6
Summary of Differences and Similarities of Evaluation Practice

Differences Among Styles	Similarities	No Pattern of Differences
1. typical day	1. time of entrance	1. development of evaluation plans
2. theoretical influences	2. purposes of evaluation	2. communication of reports
3. role of the evaluator	3. sharing of draft reports	3. ethical dilemmas
4. role of clients	4. recommendations	4. analysis of case studies
5. role of stakeholders	5. use of data	
6. when data analyses is planned		
7. types of analyses		

Epistemic Style as a Foundation for Evaluation Practice

Research Question #2

The second research question of this study was, “Does the evaluator’s epistemic style provide a foundation for evaluation practice”? If there are noticeable differences in evaluation practice among evaluators with varying epistemic styles, a typology of evaluation practice can be extrapolated. This typology would be based on epistemic style.

Epistemic Typology of Evaluators

A typology is defined as a “study of or analysis or classification based on types or categories” (Merriam-Webster’s Collegiate Dictionary, 1997, p. 1279). In this study, evaluators’ epistemic styles were measured by the Psycho-Epistemological Profile (Royce, Mos, & Learsley, 1975). The analysis of interviews of nine evaluators presented in the previous section of this chapter indicated that there were noticeable differences in practice between evaluators with varying epistemic styles. From this information it is possible to describe a typology of evaluation practice based on epistemic style, which answers the second question: “Does the evaluator’s epistemic style provide a foundation for evaluation practice?”. This typology could be useful to evaluators as a heuristic for determining or reflecting upon their own epistemic style. Because the typology is developed from a very small sample and from my personal metaphoric perspective, caution should be used in interpretation.

To create the typologies, I began with a domain analysis. Through successive rereading of the transcripts, I identified four themes: (1) theories of action, (2) attitude toward data, (3) focus on process and outcomes, and (4) relationship of theory to practice. Transcripts were then coded according to these themes.

Definition of Themes Within the Typology

Theory of Action

Argyris and Schon wrote about theories of action and their relationship to practice (1974). They describe theories as “vehicles for explanation, and prediction, and control” (Argyris & Schon, 1974, p. 5). Theories of action enable the evaluator

determine the best course of action to take when working in evaluation. Theories of action take the following form: if situation A, and you want to achieve consequence B, do C. An evaluator's theory of action will be normative for them; it will state what they ought to do if they wish to achieve their purpose.

Attitudes Toward Data

Attitudes toward data determine what evidence the evaluator considers valid and the degree of trust connected to data. According to Kasten and Ashbaugh (1991), school administrators measure all plans and actions against what they believe is desirable, or worth, important, or useful. If this also applies to the professional evaluator, personal values are going to be determinants in decisions about what problems they should focus on, which data they consider valid, and what course of action they would consider to be acceptable and plausible. This supports the radical constructivist perspective which views facts like values "in that it is a matter of individual choice as to the facts we choose to believe and the weight we give to them in our reasoning. In a sense, individuals construct their own worlds, their own realities" (House and Howe, 1999, p. xiv). Radical constructivism focuses more on the individual knower and acts of cognition rather than on social process and interaction which is known as social constructivism (Swandt, 1997b).

Focus on Process or Outcome

This theme describes whether evaluators place a focus on the process organizations use in achieving desired outcomes or whether they focus primarily on the achieved outcomes. It also looks at the link between theory, process, and outcome. This theme is grounded in the logic model that is described in the W. K.

Kellogg Foundation Evaluation Handbook (1998). This is also rooted in the work of the Swedish scholar, Evert Vedung. In his discussion of evaluation models, he explicates both impact assessment and process evaluation (Vedung, 1997).

Theory and Practice

The basic premise of this study examined the relationship of theory to practice. In this study, theory was defined as collections of models upon which action is based. This theme of “theory and practice”, refers to the type of theories evaluators choose and the way evaluators use theory in their practice. Argyris and Schon maintained that awareness of the connections between theories and practice has a relationship to effective process but lamented that this is not common practice.

Integrating thought with action effectively has plagued philosophers, frustrated social scientists, and eluded professional practitioners for years. It is one of the most prevalent and least understood problems of our age. (Argyris & Schon, 1974, p.5)

The Typology

Metaphorics

Three women evaluators representing the Metaphoric style were selected for interviews. Metaphoric B and C both worked in University settings and both had over ten years experience in the field of evaluation. Metaphoric A worked for a non-profit foundation and had less than five years evaluation experience.

Theory of Action. The Metaphorics are change agents. These are the people I picture marching in a protest line, acting out of passion rather than reason. Metaphorics are perhaps Deweyan pragmatists. Deweyan pragmatists are described as believing that “the only luck is creative action that *intervenes* in events and

transforms them into our ideal desires” (Garrison, 1999, p.319). Metaphoric B indicated acting as a change agent is her responsibility but interpreted that in more of a facilitative way.

...so you get the perspectives of the people who are trying to make the project work. I guess I’m struck by the fact that the evaluator is always an outsider and that our role should be to help those people who are trying to make something happens in schools so that seems to be very important. (Metaphoric B, p. 3)

I think of myself as a patron of schools and I want to assist, from a distance, change happening. (Metaphoric B, p.6)

Metaphoric A unequivocally chose change agent as the role she would like to play.

I started with change agent. And it’s a change agent toward improvement. I’m a big believer that we can learn things and we can improve, and that’s a real opportunity through evaluation. (Metaphoric A, p. 9)

Metaphoric C advocated evaluators should be transformational leaders. She sees evaluators as being change agents that help organizations become learning organizations that have the capacity to use evaluation for change and improvement.

...in a sense empowering people to take charge and change things. (Metaphoric C, p.11)

I am a collaborator in the process, in helping them move from one point and in a sense I am a change agent, working towards change. But not for change sake only but for the betterment of the program. (Metaphoric C, p.11)

I would like them to learn about evaluation so they can do it for themselves. To become a learning organization. It will become part of their everyday life to ask those questions. It will not be something I only come in for. That will become the culture of the organization to do some of those sort of things. (Metaphoric C, p. 12)

The metaphorics’ theory of action seems to meet some of Bolman and Deal’s characteristics of the Symbolic Leader: “inspires others to do their best, communicates a strong vision, raises enthusiasm, leads with an emphasis on culture,

highly imaginative and creative, and generates new, exciting possibilities” (1993, p. 30). Metaphoric C’s emphasis on transformational leadership fits into this model of a leader.

The transforming leader looks for potential motives in followers, seeks to satisfy higher needs and engages the full person of the follower. The result of transforming leadership is a relationship of mutual stimulation and elevation that converts followers into leaders and may convert leaders into moral agents. (Burns, 1978, p.4)

Metaphoric B also showed evidence of a symbolic leader. She talked repeatedly about her passion for education and for helping people to do their best. She hopes to awaken that passion for education and evaluation in the clients and stakeholders of evaluation. She expressed this when she responded to the questions about how she would like to involve stakeholders and clients.

I would like them to be involved early, passionately, etc. (Metaphoric B, p.12)

Actively, early, often, thoughtfully. I wish that they saw the value of evaluation always instead of incidentally, and feeling they have been forced into it by a funder who says you must have an evaluation. (Metaphoric B, p.11)

Metaphoric A also hopes to “awaken” those she works with. She chose the word interpreter to describe herself, and sees her role as not developing knowledge but as “introducing possibilities, trying new things, and asking new questions.”

Attitude Toward Data. In Royce’s book, *The Encapsulated Man* (1964), intuitionists were described as individuals who know by immediate or obvious apprehension. Intuitive knowledge is primarily conveyed via the symbol rather than the sign. This means that symbolic knowledge can convey multiple meanings, and intuitive awareness is more ambiguous than sign knowledge. This is at the foundation of the Metaphoric’s approach to data. Metaphorics rely heavily on

intuition to guide them both in collecting and interpreting data. Metaphoric A talked considerably about using intuition as the foundation of her work.

There is just a whole lot of intuitive and over time learned observational stuff. The theoretical stuff and the formal stuff are more mysterious to me and I feel that I have to modify them in order for me to accept them as practical. The other thing I could say about that is that I really struggle with what level of what kind of documentation is important in what circumstance. So in the kind of work that I do, that what I do works pretty well. It isn't always wildly legitimate to the third party. And so I need to appropriately add so that I can make my case. (Metaphoric A, p. 13).

As this comment reveals, Metaphoric A sees the need to include more formal, recognized forms of analyzing data, but generally to lend legitimacy to the analysis rather than meaning. Data is only important to the metaphoric if it lends meaning to understanding a program.

One of my frustrations at the moment is how to usefully add the more legitimate information gathering methods and to do that in a practical way. (Metaphoric A, p.13)

Standardized instruments and surveys are sort of more of the same (making general meaning) but they are even less focused on what's going on with the process in many cases. They are bits of things out of context. And I can't make much meaning from it. In a worst case scenario, they appeal to a reader that wants numbers and I don't like that. I'd rather they trust people's voices. (Metaphoric B, p. 14)

The two I prefer are descriptive case studies and content analysis. Because I think that is rooted in not only my perception of something that's useful but it's rooted in the perceptions of interviewees as near as I can determine what they are. (Metaphoric B, p.15)

Evaluation is more than measurement. If you only look at evaluation as measurement you lose too much. (Metaphoric B, p.11)

Data is only important if it contributes to meaning. "Instead of ascending a hierarchy of ever-more general criteria, all we really need is to *create* webs of better understanding" (Garrison, 1999, p. 322). The metaphors seem to be primarily meaning makers. They value people's stories and tend to be natural story tellers

themselves, using any medium that will portray a message well.

I think that I would say that extrapolating case studies and the good story is something that is really important to me. (Metaphoric A, p.15)

I got some great information from them...how they told their stories about how the program made a difference in their lives. (Metaphoric C, p.13)

Metaphorics who are comfortable with ambiguity are likely to provide multiple competing narratives of events. It may not be necessary for the Metaphoric to have definitive findings as long the findings are useful for improving the effectiveness of a program. It may be that the findings will give program staff a new perspective on their activities or program theories.

And again, hearing all of this stuff and turning it into a case example that isn't any specific case example but a set of case examples rolled into one. I think I do that without having any consciousness of it. And that's something I do a lot. (Metaphoric A, p.15)

I put describer and translator together and again that ability to talk to multiple audiences seems to be important. My favorite teacher was a freshman biology class in college and he could explain anything in 47 different ways and was willing to go to the 48th if that's what you needed and that's something I really value. (Metaphoric A, p. 9-10)

Again, this is a trait of the Dewey pragmatist who would subscribe to the ideas that there are only partial perspectives from the vantage points of various participants and there are not any absolute criteria that exist antecedent to an evaluation (Garrison, 1999).

Process and Outcome. When involved in evaluating programs, the Metaphoric focuses on the processes and then tries to draw linkages to theory and outcomes. Metaphoric B gave a good example of this characteristic when she talked about one of her theoretical influences.

I think Robert Stake was an early major influence because he talked about the missing step from what I had seen evaluation before as

being. You have some objectives, you come in and you do an assessment and there it is—the judgment. He talked about congruence. Did the activity or what people were doing match what they said they were going to do? So that was an important step. (Metaphoric C, p.3)

She spoke of the analysis of processes used in a program as both a missing and an important step in evaluation. This shows a focus on process. In discussing schools' bias toward activity, she again talked of the importance of reflecting on process.

You just have to keep doing, doing, doing, instead of thinking about how this builds on what's gone on before. What does this mean for what will come next? (Metaphoric B, p.11)

The Metaphorics also placed an emphasis on making observations in classrooms and other contexts to gather formative information to make the program more effective.

Using ethnographic methods, because what I value in classrooms is not well captured by the published observation instruments that I have seen. Again part of that is the direct quotations, descriptions of the activities that are going on and questions of what that means as opposed to forcing observations into pre-existing categories. (Metaphoric B, p.14)

Helping people understand the processes (or activities) they are using is also critical to the Metaphoric.

I think the best thing is to help people understand what they're doing, and doing it better rather than looking at someone else's and stacking up or not stacking up. (Metaphoric B, p.16)

I will go observe what they do and talk to some of the participants and see what they have learned through the process. I always like to see what is happening on the ground. (Metaphoric C, p.14)

Metaphoric B also demonstrated a focus on process when she described how she would develop an evaluation plan for a case study that was presented to her in the

interview. This case study involved the evaluation of six university grant projects that were developing instructional materials for early childhood mathematics. She presented a set of specific evaluation questions that focused on process but also inquired into the theories that were connected to the processes.

1. What materials development process(es) do the universities propose?
2. To the extent that they propose similar processes, what variables affect effectiveness?
3. To what extent do these processes incorporate the knowledge base of instructional materials development?
4. To what extent do these processes address needs of various stakeholders, e.g., pre-K students, pre-K teachers, the field of early childhood in general?

Theory and Practice. Metaphorics struggle to find a fit between theory, process, and outcome. For the metaphoric, grinding out evaluations where theory is a given and outcomes are clear is probably boring. The Metaphoric may be looking for opportunities to help define measurable outcomes, then find evaluation theories suited to the project.

Then the theorist card comes back to the secondary delight I get out of this kind of work, the intellectual fitting in frames, it's the ability to root or ground a project which is peculiar and unique in a setting which is more general. (Metaphoric B, p.10)

David Fetterman and Michael Patton are both considered influences on practice for Metaphoric B and C. It could be that this is reflective of a desire to involve stakeholders in program evaluation. But it could be evidence of this process of fitting a program into theoretical frames. Michael Patton's Utilization-Focused model of evaluation examines linkages and connections between inputs, activities, outcomes, and ultimate impacts (1997).

The problem with pure outcomes evaluation is that the results give decision makers little information to guide action. Simply learning that outcomes are high or low doesn't tell decision makers much about what to do. They also need to understand the nature of the program. (Patton, 1997, p.199)

These linkages and connections are continually being examined in Fetterman's Empowerment Evaluation model. This model is a collaborative approach involving stakeholders. Theories come from the stakeholders and are subject to a cyclical process of reflection and self evaluation. Program participants learn to continually assess their progress toward self-determined goals and to reshape their plans and strategies according to this assessment (Fetterman, 1996). Although Metaphoric A does not specify how she used theories in practice, she described a process that fit with Patton and Fetterman.

Ask good questions, collect the right information, figure out whether you learned anything and start the process over again. And that is what we call evaluation for learning is probably the out-loud framework that I have. (Metaphoric A, p.3)

The Metaphoric isn't necessarily grounded in any specific process or theory, although he/she will have favorites. The critical issue is finding the fit between theory, process, and outcomes.

Summary of the Metaphoric Style. The Metaphoric's theory of action is acting as a change agent. An idealist, the Metaphoric stirs others to action. The Metaphoric does not need to find definitive answers but is comfortable with ambiguity and multiple endings. The goal is always improvement. In gathering data, she uses her intuition to help her gather the "whole" story. Her focus is on making meaning and data are only important if they contribute to meaning. The Metaphoric tends to focus on process but is continually searching for links between theory, process, and outcome.

Rationals

Three women evaluators representing the Rational style were interviewed. Rational A worked in a public service agency, Rational B worked in a non-profit service agency, and Rational C worked in a community college setting. Rational A and Rational C had over five years experience in evaluation and Rational B had less than two years.

Theory of Action. The Rationals' theory of action is to educate. They need to have people understand their data, but just as important they feel it is their responsibility to help people understand the need for evaluation data. All of the Rationals in this study strongly expressed their desire to educate.

I think it's educating the players involved...staff, board,...educating them as to the benefits and not only of the importance of evaluation...the benefits to the agency for planning strategic planning, but also educating the staff on how evaluation can help you do your job. (Rational B, p.6)

Once you get them working on it..it's really things they can do and I like to be able to do that. To give them power to think about their research and answer their own research questions. (Rational C, p.8)

Looking at the data, looking at any surveys, continue to elaborate or describe what was said,...just trying to describe the results so people understand. (Rational A, p. 5)

Changing programs or society is important to the Rational but they prefer to take a somewhat less active part than the Metaphoric. The Rationals indicated that once they have done their job as an educator, change should be possible.

Translating the message, "here are what the results were" and translating how

somebody else has to take it and make the necessary changes. I evaluated it, I gave the data, I did the report. And how can I translate what could be some possible next steps. (Rational B, p. 6)

Just by doing the evaluation and the follow up to it, whether it's doing a report or verbal or written, that hopefully can result in change if needed. (Rational A, p.5)

You can't motivate people to change unless they want to. You can point them in the right direction. (Rational A, p. 12)

Attitudes Toward Data. According to Royce (1964), Rationals determine whether something is logical or illogical as their truth criteria. House and Howe described an approach that seems to be consistent with the Rationals in this study.

Although science and truth can be corrupted—"distorted"—by underlying values, interests, and power, truth claims are nonetheless subject to rational examination and are redeemable if generated in a way consistent with the procedural requirements of impartiality so as to prevent self-serving values the process itself can lead itself to change and what direction that change should take., interests and power from dominating. (House and Howe, 1999, p.56-57).

Rational A revealed her commitment to logic when she described how she viewed her work.

It can be fun, it can be frustrating. For a person that deals with numbers and tries to be logical, it's hard to believe that other people can look at it and not consider some of it important. (Rational A, p. 12).

Rationals trusted numbers and statistics but recognized they can be slanted for political reasons. Rational A and C preferred to use descriptive statistics, while Rational A preferred inferential statistics paired with descriptive statistics. The study of relationships through chi-squares and correlations was ranked as equally important to Rational A.

What statistics do is educate. (Rational A, p.5)

There is power in numbers and you can use it for position but to me that is the biggest misuse of data is to have an agenda to get your own thing done based on how you interpret those numbers. That's a misuse. (Rational A, p.5)

The Rationals' approach to problems is through careful analysis. Rational A maintains large databases that allow her to research a variety of questions. Rational C was concerned about evaluators using careless evaluation methodology.

They're not using proper procedures or are satisfying the assumptions of different statistical tests etc. and they're out there with this data that's really not portraying the situation as it should or not portraying the situation as well as they should. It becomes again political. (Rational C, p.13)

This supports a model of evaluator described by Carol Weiss (1998, 1972).

She puts her trust in methodology. With appropriate research tools and techniques, she expects to develop comprehensive answers to the questions with which she began the study..... The evaluator knows that absolute neutrality is impossible, but she tries to uphold the conventions of scientific research, with special emphasis on good data, sound analysis, and candid reporting (Weiss, 1998, 1972, p.98).

Focus on Process and Outcomes. Rationals have a tendency to focus on the outcomes. Rational A expressed frustration about not being able to measure some outcomes (too many intervening variables) and would like to use more inferential statistics.

How many people are impacted by what you do...you never know. It's the ripple. (Rational A, p. 3)

I'd like to see more of this done (inferential statistics). Because this really goes back to "does your program have a general outcome or are you just affecting one little segment." (Rational A, p. 11)

Rational B reinforced the frustration of sometimes never being able to assess the impact of programming.

I mean have they gained anything? Is there something someday that they'll remember that has made a difference in the way they make decisions about themselves. Self-esteem or self-reliance. It's really hard to know...but what did she learn? I find that very hard. (Rational B, p.12-13)

Rational C had a bit different outlook on outcomes. She tried to have a

continuous dialogue with clients about formative data outcomes in order to improve programming.

I usually try to involve them in the complete process...after doing the data collection you come back with some analysis and try to start discussing with them. That's the first way. But eventually you get to a report. Everyone wants a report, you have to do it. But I try to do it in a dialogue first. Then see if there are some more questions and do some more runs if there are more questions and do a little more analysis. Then we finalize it and put it in a report. (Rational C, p.12).

Relationship of Theory to Practice. Rationals appear to be task-oriented. The role of the evaluator is contingent on what type of organization they work in and specifically on what job they are called to do. They are not necessarily theory-driven.

I guess that even though I feel strongly about using theories to guide practice, I think that the idea is relative to the evaluation in working with programs you need to meet their needs and that means if it doesn't fit the theory but is still effective, that's more important. (Rational C, p. 13)

The structural frame seems to be important, meaning that productivity is emphasized, goals and roles identified, and efforts between individuals are well coordinated. Rational C's description of the stages she goes through in developing an evaluation plan illustrates this frame.

I would generally go out and meet with the clients to find out what the questions are....Then you develop a plan and then you go back and make sure the plan fits their needs...Then go through the process of developing the plan—"how do we implement it?" Then implement the plan and collect the data, analyze the data...include clients, particularly the xx program. They were involved in each step of the process. It was really participant centered. They became participants of the whole research process. So they did a lot of the data analysis, too. Not the actual running of data like SPSS, but "what does the data mean?" Then you do a report, everyone wants a report even though don't always use it. Then debrief with the client to see if any more questions need to be answered. (Rational C, p.3)

Rationals tend to operate only from proven theory and need to define objectives for their work. Theory, for them, comes from experience or is substantiated by reliable sources (probably reliable sources who have applied the theory and found it useful). There is a direct, logical connection between theory and outcome for the Rationals. When asked about theoretical influences on their work; Rational A identified experience, Rational B could not identify influential theories or experiences, and Rational C identified psychological theorists such as Piaget and Maslow. Jean Piaget believed in giving the most objective exposition possible of outcomes and leaving out findings gained from intuition (Piaget, 1952, 1963). Maslow contended that there are at least some values that do exist 'out there' in the external world, independent of human judgment (Lowry, 1973). Rational C has incorporated this into her preference for methods.

I like to use more rigorous methods (statistically rigorous). I usually use inferential statistics. (Rational C, p.10-11)

She qualified this answer though by saying that qualitative methods are more "fun." In doing evaluation, Rational C looked for theories (sometimes through searching literature) that support the perspectives of a client. This reflects a process of discovering theory that is inherent in the program development. Knowledge of the program theory will help the evaluator determine what data to collect, and what outcomes to expect. When outcomes aren't consistent with expectations, it gives her a systematic way (by examining inputs and processes) to discover possible reasons outcomes were different (Weiss, 1998). This is consistent with the Rationals' need for structure and defined outcomes.

It is not likely that a Rational will begin a project without a defined outcome and theory to support the work. Rationals generally act based on proven information and defined goals.

Summary of the Rational

The Rational's theory of action is acting as an educator. Rationals use logic as their ultimate criteria and their approach to problems is through careful analysis. Data is trusted by the Rational, as long as it has been developed from reliable and valid methodology. Statistics are preferred methods for analyzing data. Rationals have a tendency to focus on outcomes and impacts of the programs they are evaluating. They tend to operate from proven theory and need to define objectives for their work. Structural frames are important and there generally is a direct, logical connection between theories and outcomes.

Eclectics

Three women evaluators representing the Eclectic style were interviewed. Eclectic A worked in a K-12 Public School District and had over ten years of evaluation experience. Eclectic B worked as an evaluation consultant for a non-profit foundation and had 5-10 years evaluation experience and Eclectic C worked for a university and had less than three years evaluation experience.

Theory of Action. Eclectics are facilitators. They do not want to take a leading role as a change agent, but they want to facilitate the process. They are both the coach and the team mate working for a common goal. Eclectic A talked repeatedly about her desire to work with clients towards program improvement. Even in large evaluation projects, she preferred to have direct contact with people in the program.

I like working with building staff for instance...evaluating a program that resides there makes a lot of sense to me. On the other hand there are programs that go across the district...and in that case the client would be the district but I could see working with groups of people

who are involved in that, staff, teachers, parents, maybe even a core group. (Eclectic A, p.4)

Eclectic C linked facilitating along with collaborating to try to bring about change or help in decision-making. Eclectic B, who is an evaluation consultant to organizations conducting their own evaluations, saw herself as a resource. She described her organization's philosophy to program evaluation which she subscribed to.

If any additional resources were needed... we would supply them with articles, books, or even stipends to go to conferences to develop capacity. So we would try to supply them with as much knowledge and learning as possible. (Eclectic C, p.5)

Eclectics would likely agree with Egon Guba and perspectives of the radical constructivists that support the neutral role of facilitators:

Hence, evaluators should assume the role of neutral facilitators in how they handle the facts and values of other people. In this view, expertise has no special role to play. People must decide for themselves. All is relative to the individual's view (House and Howe, 1999, p. xiv).

Eclectic A strongly believed that clients should come up with their own recommendations when she first began evaluation work but has now modified her position to a process of facilitating those recommendations.

...especially when I first started evaluation, I thought the people in the program should come up with their own recommendations. If I was able to lead them to a point where they understood about the program, they could come up with the recommendations. All I had to do was put the issues on the table. (Eclectic A, p.9)

She reinforced this attitude when she discussed ethical dilemmas.

There are certainly times when evaluators have to tread very carefully because there is a fine line between a truth and a half-truth and saying things to bring about positive change and bringing about destructive consequences through evaluation. I think that's why I take the role of coach and educator and I like that better than the idea of power person or change agent...that's risky business. (Eclectic A, p.10)

Attitudes Toward Data. It was difficult to discern a pattern of attitudes toward data held across the Eclectic style. Possibly this represented the fact that this category does not have a dominant epistemic style. As epistemic style signifies how one determines knowledge, it seems logical that not having a dominant epistemic style would be an indicator of ambivalence towards data. This finding contributes to the validity of the instrument used in this study for identifying epistemic style. Eclectic A was comfortable with just about any method of data collection and analysis, although she prefers not to do case studies. Eclectic B and C prefer qualitative methods however, Eclectic C found it difficult to do case studies.

Process and Outcomes. Eclectics appear to manage process and outcome in a logical way to facilitate decision-making in the organization. This fits into the management-oriented evaluation approaches described by Worthen and Sanders.

This approach clarifies who will use the evaluation results, how they will use them, and what aspect(s) of the system they are making decisions about. The decision-maker is always the audience to whom a management-oriented evaluation is directed, and the decision-maker's concerns, informational needs, and criteria for effectiveness guide the direction of the study (Worthen & Sanders, 1987, p. 77)

Eclectic B and C defined their primary audience as those who are in positions to make decisions regarding programming.

Well, the audiences are the people in our organization who make funding decisions. (Eclectic B, p.4)

The primary audience would be administrators and funders because I would want to be using that information to change the program to meet the needs. (Eclectic C, p.4)

Outcomes are not necessarily the main objective for the Eclectics. They are more concerned with facilitating the development of good questions and collecting good information that allows the decision makers to make judgments about

programming issues.

First of all defining the questions we want to answer. Finding if the answers to those questions currently exist and if they are just being asked a different way. An assessment of what data is already there. Also finding out, exploring the intended use of the answers would be to make sure that we're asking the right questions. (Eclectic B, p.3)

For Eclectic B, finding out the intended use of the answers to the questions is an important step. This reflected a concern that the evaluation be used for decision-making. Eclectic C described an evaluation she would like to conduct which will inform decision-making in a program that she works with.

It involves evaluating whether participants in groups that we run in my other job are satisfied, whether the groups meet their needs, whether there are other things that we should offer. (Eclectic C, p.3)

Eclectic A identified the CIPP model developed by Daniel Stufflebeam (1983) as a process she used to develop questions and data collection plans. CIPP is an acronym for the types of evaluation used in the model: context, input, process, and product. Worthen and Sanders described how the CIPP model can be used as a method to develop good evaluation questions to inform decision-making (1987). Although the CIPP model was only specifically referenced by Eclectic A, it seems to represent the core approach taken by all three Eclectics. It is a way method of managing process and product to improve programming, the ultimate goal for the Eclectic.

In my mind, the most important priority is to determine the impact of the program to the participant and that goes hand-in-hand with enabling staff to improve the program effectiveness which is actually a lot of what is involved in facilitating program management. In other words, if program management is about making sure the program is going along successfully, doing what you want to do, then you need to know whether it's providing for the participants what you need to provide or what they want. (Eclectic C, p.4)

Theory and Practice. In talking about theoretical influences on their practice,

Eclectics discussed theorists or theories they have had direct experience with.

Eclectic A talked about theorists she had actually met and sometimes worked with.

Eclectic B discussed theories that were used in her organization, and Eclectic C cited a theorist whose work was the foundation of a graduate school research project she participated in. This seems to be more of a Rational approach to theory. The difference is that the Eclectics tried to make the connections between theory, practice, and outcomes like the Metaphoric.

You have to look at their planning. And if you're in on it in the beginning, I think evaluators can bring in some skills that will help do that and make good goals and measurable objectives and tie the thing together into a logic model so I think that's actually a better model.
(Eclectic A, p.2)

The difference from the Metaphorics may be more of a focus on outcomes rather than process. Eclectic A and C mention they prefer to use logic models. In fact, Eclectic A might develop a logic model to inform programming rather than a traditional evaluation report.

I have to produce something in writing...so people have a product. Somehow then it is complete. Even if it's not an evaluation report, even a logic model...something that is produced that you can take back to them and hopefully a way to talk them about the results.
(Eclectic A, p. 8)

Logic models are a picture of how the program works (W. K. Kellogg Foundation, 1998). It is a diagram that depicts the interrelationships of theories and outcomes. It identifies both the theories and assumptions underlying the program. It allows staff to develop key interim outcomes that will inform whether the long-term outcome can be expected. If the interim outcomes do not seem to match the original intent of the program, staff can go back to the original theories and revise theories or processes so the long-term outcomes are more likely to occur. Use of logic models has an implication that there is a focus on outcome. As she was sorting cards about

the purpose of evaluation, Eclectic A discussed this process of looking at what actually happens and thinking about whether it matches theories.

This one I'd like better if it said determine program impact. I'm looking for something that tells what actually happens, then the improvement thing (to enable program staff to make changes that improve program effectiveness), then how well it was thought through. I don't even care if it fits the theories but rather they even thought about, even checking their theories. (Eclectic A, p. 2)

Summary of the Eclectic. The Eclectics' theory of action is acting as a facilitator for program improvement. Eclectics, not having a dominant epistemic style, do not have a pattern of similarity in their attitudes toward data. It perhaps is based on their experience with particular methodologies or the organizational culture in which they work. Eclectics manage process and outcomes to facilitate decision-making. The Eclectics' use of theory in practice is similar to the Metaphorics' search for a fit between theory, practice, and outcomes but places more emphasis on outcomes rather than process.

Summary of Analysis

Differences in Evaluation Practice

Differences in evaluation practice were found in seven categories of analysis between evaluators representing three different epistemic styles: metaphoric, rational, and eclectic (those who did not show a dominant style). These seven categories were (1) typical day, (2) theoretical influences, (3) role of the evaluator, (4) role of clients, (5) role of stakeholders, (6) planning data analysis, and (7) types of analyses. These are important categories for understanding the evaluator and the types of evaluations they would likely conduct. Theoretical influences may give us a key to understanding the evaluators' assumptions about evaluation and perhaps program theory. The

differences between typical days give clues to how the evaluator organizes her work. The roles of evaluator, clients, and stakeholders give us definite expectations about how the evaluator will conduct the evaluation and involve clients and stakeholders. The planning of analysis and types of preferred analysis indicate what type of data we can expect. These differences are illustrated below in Table 7.

Table 7
Differences in Evaluation Practice by Epistemic Style

	Metaphoric	Eclectic	Rational
Typical Day	Variety Atypical	Exploring Task-oriented	Task-oriented
Theoretical Influences	Models that linked theories, practice, and outcomes	Theory with which they have had personal experience	Experience Well-established theory
Role of the Evaluator	Change Agent	Facilitator	Educator
Role of the Client	Active	No agreement.	No agreement.
Role of the Stakeholder	Active	No agreement.	Little, but would like more
Planning Data Analysis	Evolves throughout evaluation project	Evolves throughout evaluation project	No agreement.
Types of Analyses	Case Studies (most preferred) then other qualitative methods	No agreement but all preferred not to do case studies	No agreement but all preferred not to do case studies

A Typology of Evaluation Practice

From the analysis of interview data a typology of evaluation practice was built upon the foundation of epistemic styles. A profile of practice was developed for evaluators from the metaphoric, rational, and eclectic styles. The profiles were developed from four themes threaded through the interviews: (1) theory of action, (2) attitude toward data, (3) process and outcomes, and (4) theory and practice. The developed typology is illustrated below in Table 8.

Table 8
Typology of Evaluation Practice by Epistemic Style

	Metaphoric	Eclectic	Rational
Theory of Action	Acting as change agent	Acting as facilitator	Acting as educator
Attitude Towards Data	Intuitive Meaning-making Ambiguous	No pattern of attitudes	Logical, careful analysis Trusts numbers and statistics
Process and Outcomes	Focus on Process	Managing Process and Outcomes	Focus on Outcomes
Theory and Practice	Links theory, process, and outcomes	Links theory, process, and outcomes with a focus on outcomes	Experience Well-proven theory

From this analysis, it is possible to see a profile of practice for each of the different epistemic styles. The theory of action for the metaphorics is to act as a

change agent while the Eclectics prefer to act as facilitators and the Rationals prefer acting as educators. Metaphorics and Rationals have very different attitude towards data. Metaphorics are intuitive, interested in data only as it contributes to meaning, and are comfortable with ambiguous findings. Rationals prefer logical, careful analysis and trust numbers and statistics. The Eclectics did not show a pattern of attitudes toward data. In terms of process and outcome, Metaphorics and Rationals again take divergent actions. The Metaphorics focus on process while the Rationals focus on outcomes. The Eclectic does both by managing process and outcomes. The way evaluators used theory in practice was very similar between Metaphorics and Eclectics. They both link theory, process, and outcomes although Eclectics place a heavier emphasis on outcomes. The Rationals generally base their practice on their experience and on well-proven theory.

In my analysis, I did not intend to make the case that one epistemic style represented best practice. Although I did inform the reader that my preferred style of practice was metaphoric, that means the metaphoric style is my preference for my personal practice.

There are strengths and limitations of each style depending on from what perspective you are looking. In Chapter V, I will suggest ways in which knowledge of the influence of epistemic style on evaluation practice can be used to strengthen program evaluations.

CHAPTER V

SUMMARY AND DISCUSSION

Major Findings

Difference in Practice

The first research question of this study asked whether evaluators representing different epistemic styles exhibit differences in evaluation practice. This study provided evidence that they do. Differences in practices were found primarily in categories of process. These were roles that evaluators, clients, and stakeholders played; how analysis of data is implemented; and how their typical days were organized. Theoretical influences were also different among epistemic styles.

Epistemic Style as a Foundation for Evaluation Practice

The second question of this study asked whether an evaluator's epistemic style provided a foundation for evaluation practice. From the analysis of interview data, a typology of evaluation practice was extrapolated. It is evident that epistemic style does provide a foundation for evaluation practice. The typology was organized around four themes: (1) theory of action, (2) attitudes toward data, (3) focus on process and outcomes, and (4) theory and practice.

Although each of the epistemic styles operates with a different theory of action, they are quite close in terms of purpose. The Metaphorics act as change agents, the Eclectics as facilitators of improvement, and the Rationals as educators. All of these are related to a primary purpose of evaluation chosen by the three styles:

to enable program staff to make changes that improve program effectiveness.

Implications for Practice

Validity of Qualitative Research and Evaluation

The face of educational research and evaluation is changing (Haack, 1993). As more and more qualitative evaluation and research projects receive government funding, evaluators and researchers must pay closer attention to the validity of their research. Evaluators and researchers operating in any paradigm need to be reflective and to examine their own value structures as they design evaluation and research studies. This study provides evidence that theoretical orientations and epistemic styles have influence on evaluation practice regardless whether it is in the qualitative or quantitative paradigm. This is not necessarily a negative factor as the complexity of the world demands “views” from multiple perspectives. A problem of much evaluation research in education is that we do not bring sufficient differences to the questions; in other words, we are seeing with only one “eye”. In our attempt to eliminate the bias, we sometimes narrow the playing field. What we should be doing is illuminating the bias. By examining research with multiple lenses, we may achieve depth of vision.

Utility for Evaluators, Clients, and Stakeholders

The findings of this report have utility for three groups: (1) the evaluators, (2) the clients (who hire evaluators), and (3) the stakeholders. Evaluators, who are conscious of their epistemic style, will attempt to bring in perspectives of others to the evaluation. Incorporation of multiple-perspectives will be a more common element in program evaluation. The clients will be conscious of evaluators’ epistemic

style and reflect on whether it is compatible with the purpose they hold for the evaluation. For example, if the purpose of the evaluation is accreditation and the state is looking for evidence of outcomes, it is a wiser decision to find an evaluator with a rational or eclectic approach. This does not negate the importance of including multiple perspectives, it only gives the client some assurance of a particular framework for the beginning of the evaluation. The stakeholders, if they understand the perspective or perspectives from which an evaluation report was created, will better understand how to interpret it. They should look for evidence of multiple perspectives being included. For example, policymakers such as school board members should realize that an evaluation report developed from an entirely rational perspective may not give the board an understanding of the processes used to achieve the outcomes. This makes it difficult to make decisions about instructional practice and educational programs. If only student achievement scores are examined, it is difficult to ascertain whether the scores are an outcome of changes in instructional practice or whether there are other intervening variables. By incorporating metaphoric and eclectic approaches, linkages between theories, process, and outcome might be clarified.

Education of Future Evaluators

Professional evaluators come from a variety of educational backgrounds. There is no certification or core curriculum. Evaluators may come to their position from a background of experience or education in the field that is being evaluated. For example, in the field of educational evaluation, evaluators may have many years of teaching or school administration experience. In other cases, evaluators have post graduate degrees in the field of education. In these cases, evaluation methodologies

are learned on the job, in two or three research classes required for the degree they hold, or professional evaluation workshops. Evaluators may come from a field that requires intensive research such as anthropology. These evaluators may have a great deal of experience with methodologies but little experience with the context being evaluated. Other types of evaluators are those that have postgraduate degrees in evaluation, measurement, and research. These may be within a specific field such as education or health. The characteristic that evaluators from each of these diverse backgrounds hold in common is likely a lack of education in philosophy. It seems important, in the field of evaluation, to have an understanding of theories of knowledge. Yet, epistemology is not part of the education of a professional evaluator. An understanding of how personal philosophies influence practice could lead to greater awareness by evaluators to search out and include multiple perspectives in their work. The findings of this study provide a rationale for philosophy and personal reflection to be part of the education of evaluators.

Organization of the Evaluation Team

Findings of this study point to the value of including evaluators holding different epistemic styles on a single evaluation team. This would mean that evaluators need to be conscious of their styles and how their perspective can add value to an evaluation team. This also has ramifications for the culture of an evaluation agency. It is not enough to have evaluators from different backgrounds and epistemic styles on a team. The culture of the organization must be such that diversity in epistemic styles is a priority. Different ways of working and multiple interpretations must be valued. If multiple interpretations are reported to the client, more work needs to be done in helping clients and stakeholders to interpret and use

the findings of evaluation studies. Evaluators need education on how to build this type of culture as well as how to deal with multiple perspectives within one project.

Role and Nature of Evaluation

Traditionally, evaluation has played the role of determining the worth or merit of something (Scriven, 1973). Although this remains as a primary purpose, the data gathered in this study points to an overarching purpose of program improvement. When discussing evaluators' roles the three primary roles among the epistemic styles studied were change agent, educator, and facilitator of improvement. All of the evaluators interviewed dismissed critic, judge, connoisseur, and power merchant as roles they preferred to play. Helping organizations understand the process and outcomes of their programming and helping them use that information to improve is highly valued across groups.

This move to formative evaluation seems to be slightly incongruent with the focus of the current policy environment. Governmental bodies are focused on accountability and standards-based programs. High stakes are being attached to outcomes. Schools and public agencies are being punished (by withdrawing government funding) if standards are not met or outcomes are not considered acceptable (Walsh-Sarnecki & Van Moorlehem, 1998, June 8). Resources are slim for using data to improve programming. In Michigan, potential financial resources are being used to finance rewards (in the form of small scholarships) for those students who are high achievers rather than using the funds for improving schools that need help with their educational programming (Johnson, 1999, June 10). This type of policy makes the assumption that schools will improve if only a large enough carrot is held out. It fails to consider the idea that lack of achievement may be an

outcome due to lack of human or financial resources. More importantly, it fails to recognize that the standardized tests may be biased toward particular populations or may not be an accurate measure of the type of knowledge we hope our students are learning. In the words of Alfie Kohn, “raising scores is completely different from helping students to learn” (1999, pp. A31).

If evaluators are serious about wanting to be change agents, educators, and facilitators of improvement; they must take stronger stances about formative evaluation and using evaluation data. When presenting evaluation contract proposals, these elements should be embedded in the evaluation plans. In evaluation reports, they should be central to the study. In essence, evaluators must become lobbyists to policymakers to advocate the use of evaluation for program improvement. An evaluation culture should be built in order that evaluation is done not as an “after-thought” but as a process embedded into the culture of organizations, allowing for continuous renewal and improvement.

Limitations of the Study

Influence of Organizational Culture

Program evaluation can be an extremely complex process to deconstruct. Not only do the interactions of evaluators with program staff and stakeholders need to be considered, but also the organizational culture and context may erect limitations to the nature of evaluations that can be conducted (Marais, 1998). In real practice, “evaluators do not necessarily consider all data or all potential audiences, only those relevant to a particular time and place in a particular context” (House and Howe, 1999, p. xvii). This seemed to “hold true” during the interviews of evaluators. Even though evaluators were asked to respond with how they ideally would like to practice,

most referred back to how they actually do practice. This did mean that the specific organizational culture and context within which they worked, possibly influenced their responses. This study was not able to separate out the influence of organizational culture.

Small Sample

The small sample possibly limited the degree this study can be generalized to a larger population. It does show that within this sample, epistemic style (as measured by the PEP) was an influence on practice.

Espoused Views

The interview data only reflected evaluators' espoused views. It would lend validity to the study if data from other sources had been included. This would entail an ethnographic study that includes job shadowing and collection of evaluation plans and reports.

Lack of Empirical Style

The empirical style was not represented in this sample as a dominant style. Further study is needed to determine whether this style is not represented in the population of evaluators or whether this group did not care to participate in the study. As positivism is receding as the predominant paradigm in the social sciences, it is not surprising to find that the empirical style is the least dominant style in the hierarchy of evaluators' epistemic style.

Sole Author

The concluding limitation of this study is that I was the sole author. I conceptualized the structure of the study, the research questions, and the interview protocol. Carol Weiss used a quotation from Werner Heisenberg in the second edition of her book, Evaluation (1972, 1998, p. 114) that talked about this dilemma: “Since the measuring device has been constructed by the observer... we have to remember that what we observe is not nature itself but nature exposed to our method of questioning.” A positive aspect of this study is that at each step, other viewpoints were solicited. The themes I developed from the interview data, also possibly limited my vision. After all, as a Metaphoric, I was only interested in data that helped me make meaning. To counter this limitation, I have tried to make my process and bias as clear as possible, letting others judge my analysis of the data. Others may find competing narratives I have not explored.

Suggestions for Further Research

Expansion of Interview Protocol

The first recommendation I make is to extend this study. The expansion of the interview protocol to gather more information about background of interviewees would be helpful. This would include the degree of education in professional program evaluation and the types of organizations/institutions the evaluator has worked previously. By having these additional variables, an analysis could be done as to whether the patterns of practice are indicative only by epistemic style or whether other intervening variables have influence.

Larger Sample

If a larger interview sample was used that included males, it could be determined whether males show evidence of the same patterns of practice according to epistemic style. In a larger sample it would be possible to control for other variables such as type of organization in which the evaluator worked. To control for organizational culture, several individuals (representing different epistemic styles) from the same organization could be studied to determine whether the same pattern of practice according to epistemic style exists within the same organization.

Ethnographic Methods

Richer data could be gathered from an ethnographic study of each evaluator rather than the short interview used in this study. This would mean not only a longer interview, but also job shadowing them on multiple days to get a better picture of practice. The examination of evaluation plans and reports developed by the evaluators would also be useful in discerning actual practice. Of course, this is not a practical process. Resources for such work are quite limited. In addition, the nature of evaluation work sometimes requires confidentiality (particularly with regard to evaluation reports).

Educational Researchers

The influence of epistemic style is relevant to other fields than evaluation. As evaluators and researchers have a close link, it makes sense to do a similar study of educational researchers. Such a study would perhaps bring attention to the need for researchers to include their personal frameworks within their studies, allowing the users of the research (policy-makers, school administrators, teachers, families) to

judge the validity of the research for their decision-making.

School Administrators and Teachers

Extension of this study to school administrators and teachers is warranted as evaluation is a major component of the responsibilities of both groups. By examining the influence epistemic style has in developing policies and designing instruction or curricula, it would draw attention to the need for understanding one's personal theories and recognizing the influence they have on practice.

Multiple Perspectives

Finally, I would pose the research questions to researchers operating from different epistemic styles and have them develop their own study and analysis. It would be interesting and useful to discover whether different theoretical frameworks and methodologies would result in different findings for this study.

Conclusion

Link Between Epistemic Style and Evaluation Practice

It is unrealistic to divorce one's epistemic style from evaluation practice. Furthermore, when an evaluator is contracted to evaluate a program, it will always be somewhat impositional. As Daly wrote, "by virtue of our roles as researchers, observers, commentators, and social scientists, we are there to organize, select, and construct explanation." (1997, p. 350).

It is most likely a difficult proposition for many evaluators to give up the idea that they can be the ultimate (or at least reliable) source of knowledge for the programs they evaluate. It is perhaps even more difficult for evaluators to recognize

their evaluation reports may be mixed with their errors, biases, and autobiographies. The acknowledgment that their work is mixed up with their hopes and dreams is probably most difficult but essential if evaluators are to value the use of multiple perspectives in evaluation.

Utility of Multiple Perspectives

What might happen if, instead of resisting ways of thinking and doing of other evaluators, we tried to approach programs from more than one perspective in order to make sense and meaning for both program administrators and participants? If the audiences of evaluation reports more consistently understood and respected the range of assumptions shaping evaluation reports, would interpretation of the reports change? Would the reports have more utility to the user? As a metaphoric, I believe that multiple narratives construct a locus for growth. It allows the democratic process to take root in communities, institutions, and organizations. It requires the reflective process—a critical component in education, public administration, and a democratic society. Conflict can then be recognized as a resource rather than as a threat.

Theory-Laden Observations

In much of the literature, the idea of theory-laden observations has a pejorative nature. I maintain that theory-laden observations can be enabling. We often need theories to explain or make sense of what we observe. The key is to be conscious of our theories in use and to be aware of how they are influencing our observations of the world. Remembering that reality makes sense but what makes sense is not necessarily real is a critical theory to remember when we are interpreting our observations. This consciousness will be helpful in doing program evaluation. It

points to the need to solicit multiple perspectives while evaluating programs.

Practical Problems

There are some practical problems these ideas entail. “How do I know what epistemic style my evaluator represents and their profile of practice?” It is not likely evaluators will begin tattooing this information on their bodies or carrying around ID cards proclaiming their profile of practice. Our only avenue is to engage evaluators in meaningful dialogue regarding personal theories and styles of practice before an evaluation begins. The hiring of an evaluator should be approached with as much caution as hopefully is used in hiring a teacher.

The life and demise of programs often rest on the findings of evaluation reports. The continuous improvement and renewal of programs are also often the product of a formative evaluation process. The stakes are too high, particularly in the fields of health and education, to allow evaluation to be considered a mere technical process. It is more than that. It is time that it is given thoughtful, careful, and knowledgeable consideration.

The Future

The findings of this study provide evidence that personal epistemic styles (theories of knowledge) of evaluators do influence evaluation practice. This doesn't mean that our evaluation studies are invalidated, it only means they perhaps represent only one perspective. If a philosophy of multiple-perspectives was valued, this would mean that people would always be open to learn other viewpoints and possibilities. Stakeholders of an evaluation would learn to take active roles in program evaluation in order that their perspectives are represented. Clarity would possibly develop from

looking through different lenses. This state of mind would contribute to a continuous, learning approach to evaluation. Organizations would not be content with a final evaluation report. In fact, evaluation reports might no longer be relegated to file drawers but become dynamic, living documents that improve people's lives.

Appendix A
Autobiographical Bias

AUTOBIOGRAPHICAL BIAS

This study examines whether one's personal theories of knowledge have an influence on one's evaluation practice. I believe any question chosen for study is not chosen innocently. The seed for inquiry is planted in our history. To make such statements, I must examine why it is that I chose this topic for study and make that investigation open to the consumers of the research. To do otherwise, would seem to be duplicitous (at least from my perspective).

To tell my history and consider how it has formed my present practice is difficult. Not because of a traumatic past but because my memories are so rich from the time I was a very young child. I cannot possibly tell "all" in a few pages but yet "all" of it makes up my present. So I will sail through the depths and try to pull out the first memories that seem to point to the formation of the questions in this study.

I don't remember a time when I wasn't interested in books and stories. There was always someone to read them to us. Born in a small town in Northern Michigan, my childhood would seem idyllic to many. Life began in the small Finnish town of Kaleva, named after the Finnish epic poem, The Kalevala. At that time, my family consisted of my parents, two teen-age uncles, and my older sister. My family left the family homestead before I could develop memories of Kaleva as my home (I was only 2) but visits to relatives were frequent over the years and I developed a "healthy" Finnish appetite for sweet rolls, stories, and coffee.

My next home was in Harbor Springs, Michigan where we lived on a high bluff across from Lake Michigan. Our landlady lived in a big house next door, which fascinated me with its shelves and shelves of books as well as other fancy things behind glass. My sister and I were very excited when she gave us some of the books her children had outgrown and soon my sister Rita, who was 4, was reading them to

me.

My fascination with stories grew with the introduction of the public library when we moved to a small town in Southwestern Michigan. Unfortunately, on the first visit, I discovered I had to be 5 years old before I could get my own library card. Fortunately, the librarian took pity on me and said that if I could learn to sign my first and last name, I could also have a library card. Suddenly, nothing had seemed so important. It seemed forever but I don't think it was very long before I could write small enough to fit on the space on the form. This whole process is very bright in my memory—maybe because it was so important to me. Saturday mornings at the library, with its exciting story-telling hour, became a very important part of our lives.

But the excitement of getting my own library card was small compared to the prospect of going to school. I had already learned to read but hadn't yet discovered how to write my own stories. Most of all, I was interested in having lots of friends. My kindergarten experience quickly turned into something I had never imagined. First of all, I had to go to some special testing that only a few children were selected for. Then I discovered that I talked "funny" and I would have to go to special speech lessons each day. Until that point, I never knew that I didn't sound right. Suddenly it seemed like no one would play with me on the playground. My best friends became the other two kids who had to go to speech class. School was dreadful. I had no chance to show how well I could read (this was a Kindergarten of the 50's where sandbox and clay table were major activities) and my claim to fame was the girl who talked funny.

I soon discovered a strategy for getting along—I just wouldn't talk much. The teacher seemed to like the quiet kids and at least I wouldn't have to listen to her correct my speech every time I talked. I still managed to get into some trouble by not

following directions. Much of the time, I just didn't understand what she said. That brought on the hearing tests. The tests were pretty easy though, you only had to say when a sound started and stopped. I thought I would only be in more trouble if I explained that I could always hear words but sometimes they were mumble-jumble. Surely that would mean I was retarded or something so I kept that as my secret.

My school-life soon changed dramatically. In the middle of the year, we moved five miles outside of town and we transferred to a one-room school. Although it was scary at first having eighth graders in the same room as Kindergartners, I was also very excited because Kindergartners had regular reading class! We also had workbooks we could write in and we were expected to practice printing every day! It was a requirement to keep tablets to draw in and library books to read in our desk so we wouldn't disturb anyone when we were done with our work. This was everything I imagined school was like and everything that I was good at. I no longer felt strange. I knew I would have a lot of friends here.

Then I got in trouble again. First, I had to spend recess inside for blowing bubbles in my milk with a straw and making kids laugh. The teacher did not realize I didn't understand how to use the straw. Shortly after that episode, the weekly visiting music teacher was teaching us a dance that involved skipping in a circle. Soon everyone discovered I didn't know how to skip. I missed most of my recesses that week because I had to stay inside and learn to skip with the teacher. I was positive that soon everyone would know that I couldn't catch a ball, skate, or talk very well. However, this time, things were different. My parents learned I was missing recess and my skipping lessons stopped. Eventually, I learned how to do all of those things that everyone seemed to think were so important. Meanwhile I reveled in being the star reader in the class. My only bad moments were when I had to see the visiting

speech therapist once a week. I was ashamed to go, but I don't remember anyone at this school making a big deal about it

The next year when I was in first grade, I was allowed to help other kids that were having trouble with reading. I liked to rush through my work so I could work on my projects which may be reading, making paper doll clothes, making scrapbooks from old magazines, drawing, writing stories, and on and on. Arithmetic was the only drudgery. After school time was filled with playing "pioneers" in the woods, walking logs in the swamp, riding bikes down dirt roads, and playing with dolls.

Indian Lake School was a rare opportunity, I guess, for a child of the 50's. Mothers sometimes came in to show us how to do different things. Besides our "regular" subjects, we learned to bake bread, eat exotic foods like octopus, organize doll shows, and sing in foreign languages. My sister and I were proud to teach "Jesus Loves Me" in Finnish.

The reader may wonder why I digress at length about these childhood experiences. My experiences with Indian Lake School provided the foundation on which I have built my teaching and evaluation practice. My initial experience with school was framed by testing; testing that showed that I had problems. I don't question the merit of being tested—it was important to recognize my speech problems and provide me with the expertise that could correct them. It was also important to have my hearing tested but unfortunately the test was not comprehensive enough to diagnose my problem. The fact that the teacher framed her picture of me by the outcome of these tests without knowing what else I could do was problematic. In retrospect, I see there were probably not a lot of options—she had 40 other five-year-olds in the classroom to deal with.

I feel I was blessed to have attended a one-room school for most of my

elementary school experience. At a small school like Indian Lake, there wasn't really any "norm." We only had 1-4 students per grade and there were always migrant children from the farms coming and going. Instruction was very individualized. I often would have my reading lessons with older kids but never worked ahead of my grade in arithmetic. I loved art and was allowed to create it whenever my work was finished. As I think back now, it was truly a school for the imagination. It was an environment where everyone was accepted (even if they did talk and walk a bit funny). Families played a big part in the school. All parents were expected to go to school board meetings frequently. In warm weather, even the kids went and played outside while the adults inside talked. Just how would this type of school be evaluated today? In Michigan, it would likely be closed. Its curriculum was not aligned with a standardized test. Standardized test scores, I am willing to guess, had a high degree of variance. It would not get points for students being independent learners, or being skilled in both the practical and imaginative arts.

Being a seeker, it is not surprising I became a teacher. During my years in preparation for becoming a teacher, I was very excited about the "Open School" movement. One semester I volunteered at a local school that was actually using the concept of learning centers to frame their instruction. My enthusiasm was grounded in the idea that there was a way I could re-create my childhood experiences in the one-room school. Student teaching dashed some of that enthusiasm. I soon learned it was necessary to write behavioral objectives for each activity that occurred in the Kindergarten classroom where I would complete my internship. I held on to the hope that when I had my own classroom things would be different.

My own classroom was different but not particularly in ways that I expected. I taught in a rural area and generally had two grades in my elementary classroom.

Curricula were defined by the choice of textbooks we were expected to use.

Evaluation of students was largely determined by standardized tests. Evaluation of teachers seemed to have a direct correlation with how well behaved the classrooms were and the popularity of teachers with parents.

There were many merits of this school as there were some excellent teachers and involved parents. But it was difficult for me to accept that there were also some teachers who were doing a “sloppy” job and not paying any penalty. I worried about the kids who were not getting the extra help they needed. I particularly worried about children in my own classes who had great difficulty learning to read yet seemed very bright in most other ways. I felt that there must be some method of teaching that I had not learned. I felt sorrow that these students would forever be stigmatized as being slow when perhaps the reality was methods of teaching and testing their intelligence were inadequate. Although issues of equity in terms of race and gender were beginning to be addressed, equity in terms of learning differences was largely unrecognized as a problem in public schools.

In 1990, I learned about a program for training people to teach dyslexics. At the time, the tuition seemed exorbitant but I had the nagging feeling it would give me some answers about teaching people with learning differences. Thus began my association with the Michigan Dyslexia Institute. In the space here, I cannot begin to tell the rewards of teaching people to read who had come up against failure so many times before. The joys of these experiences were sadly shadowed with the experiences these students and their parents faced each day with schools, both private and public. It was a constant struggle to obtain appropriate evaluation for these students. The difficulty was that if students were diagnosed with a learning disability, the school was obligated by law to provide appropriate education. In the case of

dyslexia, schools would need to provide teachers that had specific training in working with dyslexics. Because schools had a vested interest in not finding a student to be dyslexic, the testing process was questionable. Consequently, many parents hired an outside evaluator.

After these experiences, I felt a need to become more knowledgeable about testing practices. In the fall of 1994, I began a Masters program in Educational Evaluation, Measurement, and Research. In 1996, I began working as an evaluator for a SAMPI, an evaluation unit within the Department of Science Studies at Western Michigan University. The staff at SAMPI includes 6-8 professional evaluators with varied experiences from elementary teaching to scientific research. The mission of SAMPI is to improve science and mathematics programming in public schools through evaluation. Now it may seem that our varied backgrounds would be problematic but I see that as our strength. Each of us, in my opinion, approaches classrooms from a different perspective. Yet, our evaluation plans tend to have a common pattern. Coming to a consensus on evaluation plans and on interpretation of gathered data sometimes spawns a great deal of angst among staff. There are some who want to come up with common criteria for evaluation plans as well as data gathering, interpreting, and reporting. For some projects, this seems appropriate. For instance, when data from many programs across a region or across the United States are being aggregated, it is important to know that there is a degree of commonality in how data is collected and aggregated. But my philosophy of evaluation is grounded in my experiences as a child and as a teacher. I feel it is critical to include as many perspectives as possible in the practice of evaluation. SAMPI's diversity in its staff can be its strength. In this way, the likelihood of someone becoming "invisible" or the strengths and weaknesses of a program going unrecognized because someone

didn't think to ask the question are diminished. We all give different evidence different weights. Our goals should be to gather information in as many ways as possible and to invite all into the inquiry process. If we can find a thread running through several different perspectives, perhaps we are can see a clearer picture of "reality".

It is still fair to ask, "how would I ideally like to evaluate a classroom?" The answer is easy for me although highly impractical. I would want to know the classroom story. I would want to observe how the students and the teacher interact. I would want to see evidence of imagination and compassion. I would hope to see evidence of independent learning, a classroom of seekers. Of course I would also want to see evidence of growth. This would include test scores, projects, writings, and records of other creative endeavors across the year. This is not a particularly novel approach; it is what portfolio assessments are all about. But I would want to see this evidence not just for students but for teachers as well. I realize this is kind of an idealistic stance that is highly impractical in the context of most of our evaluation contracts. But if instruction is guided by evaluation and assessment (as many claim) it is essential that we start evaluating in a way that is cognizant of the deep complexity and diversity of human behavior and cognition.

Appendix B
Human Subjects Institutional Review Board Approval



WESTERN MICHIGAN UNIVERSITY

Date: 27 August 1999

To: Jianping Shen, Principal Investigator
Deana Draz, Student Investigator for dissertation

From: Sylvia Culp, Chair *Sylvia Culp*

Re: HSIRB Project Number 99-07-19

This letter will serve as confirmation that your research project entitled "A Study of the Relationship of Epistemic Style to Evaluation Practice" 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.

Please note that you may **only** conduct this research exactly in the form it was approved. You must seek specific board approval for any changes in this project. You must also seek reapproval if the project extends beyond the termination date noted below. In addition if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

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

Approval Termination: 27 August 2000

Appendix C
Informed Consent Form

Western Michigan University
Department of Educational Studies

WESTERN MICHIGAN UNIVERSITY
H. S. I. P.
Approved for use for one year this date:
AUG 27 1999
**Sylvia Culp*
HSIRB Chair

You are invited to participate in a research project entitled "The Relationship between Epistemic Style and Evaluation Practice," being conducted by Dr. Jianping Shen and Deanna Draze from Western Michigan University, Department of Educational Studies. This research is being conducted as part of the dissertation requirements for Deanna Draze. The purpose is to examine the differences between personal theories of knowledge of professional evaluators and determine the influence these differences have on evaluation practice.

You are invited to complete an instrument that categorizes a person's epistemic beliefs (theory of knowledge). This instrument, the Psycho-Epistemological Profile is composed of 90 items and will take approximately 20-30 minutes to complete. Your replies will be completely anonymous, so do not put your name anywhere on the form. Although there are codes on the instruments, these will be removed after the data collection process. Names will not be connected to any of the data reporting. If you have any questions, you may contact Dr. Jianping Shen at (616-387-3879), Deanna Draze at (616-398-2418), the Human Subjects Institutional Review Board (616-387-8293), or the vice president for research (616-387-8298).

This consent document has been approved for use for one year by the Human Subjects Institutional Review Board as indicated by the stamped date and signature of the board chair in the upper right corner. You should not participate in this project if the corner does not have a stamped date and signature.

Appendix D
Letters to Evaluators Requesting Participation

Date

Name

Address

City, State

Dear Name

I am a doctoral student at Western Michigan University in the Department of Educational Studies' Measurement, Research, and Evaluation Program and also a Research Associate on the evaluation team at SAMPI (Science, Mathematics, and Program Improvement) at Western Michigan University. To complete my Ph.D., I am currently working on my dissertation research which focuses on the relationship between epistemic style (theory of knowledge) and evaluation practice.

The first portion of the study will be assessing the reliability of the instrument I plan to use for measuring epistemic style, The Psycho-Epistemological Survey. This instrument has been extensively used with graduate students but has not been tested on a population of professional evaluators. I would like to invite you to participate in this reliability study. It will take approximately a half-hour of your time. The instrument is enclosed with this letter. Please return it in the stamped, addressed envelope I have enclosed.

If your instrument, is one of 9 randomly chosen instruments, you will receive a request for an interview. The purpose of this interview is to study the relationship between epistemic style and evaluation practice. This is explained further in the enclosed document. The interview, if you are chosen, would take approximately one hour. Participation, of course, is entirely voluntary.

In this mailing, you will find an informed consent document and a brief abstract outlining the study, its purposes, procedures, and schedule. After reading this information, I hope you will agree to take part in the study.

I will greatly appreciate the time taken to help me in my dissertation research. In appreciation, I will send a summary of my findings to those who participate in the study. If you have any questions, I can be reached at the above phone number evenings and weekends and at 616-387-2418 during the day. If you would like your individual score, enclose a separate note with your instrument. Please do not put any identifying ID on the instrument.

Thank you for your time and consideration.

Deanna Draz

Date

Name

Address

City State Zip Code

Dear Name,

Thank you for your recent participation in my dissertation research. Through your generous volunteer of time, I have been able to assess the reliability of the instrument I have chosen to measure the epistemic style of professional evaluators.

The purpose of the second portion of the study is to examine the relationship between epistemic style (theory of knowledge) and evaluation practice. This analysis will be conducted through personal interviews of nine professional evaluators. This sample was chosen from the sample of evaluators who took part in the reliability study. After the instruments in the reliability study were scored, they were sorted into three groups representing the three epistemic styles. From each group, three instruments were randomly chosen. The codes were then matched to a participant file in order that participants could be contacted. Your instrument was one of the randomly chosen instruments. I invite you to be part of the interview portion of this study.

The interview will take approximately one hour. My goal is to elicit participants' perceptions and understandings of their evaluation practice and examine whether these descriptions of practice are similar to evaluators with the same dominant epistemic style. Topics to be addressed will include methods of developing evaluation plans, methods of data gathering, involvement of stakeholders/clients, methods of communication and/or reporting, involvement in the utilization of evaluation data, and ethical dilemmas. I will provide you with a transcript of the interview if requested. You will also be sent a draft copy of the analysis for written review and revision. I will contact you within one week, to further discuss the study, answer any questions you might have, and ask for your participation.

I will greatly appreciate the time taken to help me in my dissertation research. It is my hope to add to our profession's knowledge about evaluation practice. If you have any questions, I can be reached at the above phone number evenings and weekends and at 616-387-2418 during the day.

Thank you for your time and consideration.

Deanna Drazee

Date

Name

Address

City, State

Dear Name,

Recently you received a document requesting participation in my dissertation research which focuses on the relationship between epistemic style (theory of knowledge) and evaluation practice. I realize that you are very busy and probably haven't had time to reply. I am hoping you will still be willing to participate.

The first portion of the study will be assessing the reliability of the instrument I plan to use for measuring epistemic style, The Psycho-Epistemological Survey. This instrument has been extensively used with graduate students but has not been tested on a population of professional evaluators. I would like to invite you to participate in this reliability study. It will take approximately a half-hour of your time. For your convenience, I have enclosed another copy of the instrument with this letter. A self-addressed stamped envelope is enclosed for your response.

If your instrument, is one of 9 randomly chosen instruments, you will receive a request for an interview. The purpose of this interview is to study the relationship between epistemic style and evaluation practice. This is explained further in the enclosed document. The interview, if you are chosen, would take approximately one hour. Participation, of course, is **entirely voluntary**.

In case you do not have the original mailing available, I have again enclosed an informed consent document and a brief abstract outlining the study, its purposes, procedures, and schedule. After reading this information, I hope you will agree to take part in the study.

I will greatly appreciate the time taken to help me in my dissertation research. In appreciation, I will send a summary of my findings to those who participate in the study. If you have any questions, I can be reached at the above phone number evenings and weekends and at 616-387-2418 during the day. If you would like to receive your individual score, send a separate note with your instrument. All stored data should not have any personal identification recorded on it.

Thank you for your time and consideration.

Deanna Drazie

Appendix E
Permission to Use Psycho-Epistemological Profile



UNIVERSITY OF ALBERTA

January 21, 1999

Ms. Deanna Drazé
1006 Piccadilly Rd.
Kalamazoo, MI 49006
USA

Dear Ms. Drazé,

Thank you for your Express of letter January 16th. I did forward a copy of the PEP after our telephone conversation; it must have been lost or else I had the wrong address.

In any case, herewith another copy of the PEP.

While I continue to receive requests for the Manual at the rate of about 2-3 mos., I have not worked with the PEP since my undergraduate days in the sixties when I was Joseph Royce's research assistant. Clearly there remains considerable interests in the question of "style"; perhaps, it is time for someone to do an update on the literature and improve the instrument - in short, to write a new Manual.

I trust PEP will prove useful.

Sincerely,

Leendert P. Mos
Professor

Department of Psychology

1-220 Biological Sciences Building • University of Alberta • Edmonton • Canada • T6R 2H1

Telephone: (403) 494-3216 • Fax: (403) 492-7568

www.ualberta.ca

Appendix F
Interview Protocol

INTERVIEW PROTOCOL

1. Self-introduction and explanation of the purpose of the study
2. Explanation of the Informed Consent Form
3. Explanation of recording and transcription procedures

QUESTIONS

1. Can you describe a typical day in your work?
2. In what types of organizational settings do you normally conduct evaluations?
Are you interested in conducting evaluations in other types of settings?
3. Can you name theoretical influences on your practice? (You may wish to name individuals connected to these theories.)
4. What purposes of evaluation do you give priority? Do you feel this is compromised by your present contracts or the organization for which you work?
5. When do you most often enter the evaluation process?
 - Before program begins to assist with the needs analysis stage
 - During the program development stage.
 - When the objectives of the program have already been established.
 - After the client has developed specific questions he/she wants answered.
 - After data has been already collected.
 - At the end of a program

Which do you prefer?

6. What are the stages you go through in developing an evaluation plan?
7. I am going to show you five cards with different purposes of evaluation. Would you rank order them according to what type of evaluation you prefer to do?
Cards read:
 - To document whether program objectives have been met
 - To enable program staff to make changes that improve program effectiveness.
 - To foster accountability
 - To determine the effectiveness of programs for participants
 - To determine whether implementation and outcomes of the program “fit” with the theories the program is based on.
 - To facilitate program management

8. Who are the primary audiences for the evaluations you conduct?
- Program administrators?
 - Program participants?
 - Advisory or governing boards?
 - Program Funders ?
9. I would like you to talk about your ideal role as an evaluator. I will lay out cards with various roles. Sort out those roles you would most like to play in future evaluations. Cards read:
- | | |
|----------------|------------------|
| • Facilitator | • Critic |
| • Expert | • Consultant |
| • Investigator | • Connoisseur |
| • Patron | • Power merchant |
| • Translator | • Interpreter |
| • Theorist | • Educator |
| • Coach | • Describer |
| • Collaborator | • Negotiator |
| • Judge | • Change agent |
| • Mediator | |

Can you rank your selected cards in order to their priority for you as an ideal role you would like to play? Can you define each of the roles you have chosen?

10. When you are involved in program evaluation, what role do clients (program administrators) take? Do you involve stakeholders in an evaluation? (for example in evaluating a new instructional program at a high school, do students or parents play any active part in the evaluation?)
11. What are the primary methods you use in collecting information for your evaluation? Choose from the cards provided or add any additional methods. Cards read:
- Standardized instruments/surveys
 - Self-developed surveys
 - Document analysis
 - Focus groups
 - Classroom observations using a published observation instrument
 - Classroom observations using ethnographic methods
 - Interviews
12. At what point do you typically plan the analysis of your data?

13. What type of analyses do you typically do? Choose from the cards provided or add any additional analyses you prefer. Cards read:
- Descriptive- quantitative methods such as sums, percentages, averages
 - Descriptive- content analysis of qualitative data such as interviews, focus groups, and open-ended questions on surveys
 - Descriptive- case studies
 - Study of relationships using cross tabulations or correlations
 - Inferential statistics
 - Control groups
14. What is your typical way of communicating evaluation findings to clients and stakeholders?
15. Do you share draft versions of evaluation reports with anyone before the final report is presented?
16. Do your evaluation reports include recommendations for program improvement?
17. Do you play a part in helping organizations use evaluation data on a regular basis?
18. Is there anything else about your evaluation practice that we haven't discussed, you think is important to know?
19. Can you describe ethical dilemmas you run into in your practice that are particularly troublesome to you (please do not name any specific names or organizations)?
20. Prior to this interview, you were given a short case study to review. Did you have time to reflect on it? What type of evaluation plan would you devise for this case?

Appendix G
Case Study

CASE STUDY

Six universities have won large grants to develop instructional materials in mathematics for pre-K classrooms. You have been contracted by the funder of these grants to evaluate the instructional materials development process across the six projects. The six universities are located 500-800 miles from your location. The individual university grants extend four years. Your contract will begin six months after their grants begin and will extend three years. You are allocated \$90,000 a year to conduct the evaluation. What approach will you take? Briefly describe your evaluation design and rationale. Keep the budget in mind, but there is no need to concentrate on detailed costs.

Appendix H

Table 9: Interview Responses

Table 9
Interview Responses

Categories	Metaphoric	Epistemic Styles	
		Eclectic	Rational
Typical Day	<p>A. Typical day is atypical</p> <p>B. Multiple projects going on</p> <p>C. Not only doing evaluation</p>	<p>A. Web search, grant development, meetings, conferences, working on research questions</p> <p>B. Developing tools, looking at instruments, meetings, consulting, funding decisions</p> <p>C. Looking for relevant information, focusing on evaluation data, making notes before writing</p>	<p>A. Collecting data, managing data-bases, generating reports</p> <p>B. Grant developing process</p> <p>C. Transitioning from evaluation to assessment, getting facts, selecting instruments, determining analyses, methodologies, providing support</p>
Theoretical Influences	<p>A. Based on school board experiences, outcomes based evaluation, and the concept of evaluation for learning</p> <p>B. Stake's congruence model, Fetterman's empowerment evaluation</p> <p>C. Patton's Utilization focused, Fetterman's Empowerment, Learning organization/organizational change models, outcome based models</p>	<p>A. Stufflebeam's CIPP model, asking good questions, Patton, Guba-getting strong qualitative data</p> <p>B. Wholey's Logic Model, Patton, Organizational models through the independent sector</p> <p>C. Naturalistic models, Guba</p>	<p>A. Years being a teacher, understanding how people approach knowledge and convey that to people</p> <p>B. No</p> <p>C. Undergraduate and graduate experience in educational psychology program, Piaget, Maslow, looks for theories that support perspectives of participants</p>

Table 9— Continued

Categories	Epistemic Styles		
	Metaphoric	Eclectic	Rational
Preferred Entrance	A. Before B. Before C. —	A. Early B. Before C. Before	A. Depends B. After C. Before
Development of Evaluation Plan	A. Problem solving method, define the problem B. Collaborative, reflect on the problem, discuss indicators and measures, look at prior knowledge, collect data C. Collaborative, reflect and define questions, study documents, collect data	A. Develop questions. Use CIPP model, create tables, discuss with clients B. Define questions, discuss where answers may be, discuss intended use, supply organization with as much knowledge and learning as possible C. Start with focus groups to define questions, design survey	A. Depends on context B. Using focus groups for developing questions C. Meet with clients, define questions, develop plan, present plan to clients, match needs, implement, collect data (include clients), interpret data, report, debrief
Purposes of Evaluation	A. (1) Enable program staff to make changes to improve program effectiveness (2) Determine impact on participants (3) Determine outcomes fit theories (4) Determine whether objectives have been met (5) Facilitating program management (6) Fostering Accountability	A. (1) Enable program staff to make changes to improve program effectiveness (2) Determine impact on participants (3) Determine outcomes fit theories LAST (facilitating program management, determining whether program objectives have been met, fostering accountability)	A. (1) Needs assessment (2) Determine outcomes fit theories (3) Enable program staff to make changes to improve program effectiveness LAST: Facilitating program management and fostering accountability

Table 9—Continued

Categories	Epistemic Styles		
	Metaphoric	Eclectic	Rational
Purposes of Evaluation (cont.)	<p>B. (1) Determining how program is impacting participants (2) Enable Program staff to make changes to improve program effectiveness (3) Foster accountability (4) Determining whether outcomes fit theories (5) Document whether program objectives have been met (6) Facilitate program management.</p>	<p>B. (1) Enable Program staff to make changes to improve program effectiveness (2) Determining how program is impacting participants (3) Facilitating program management, determining whether program objectives have been met, fostering accountability LAST: Determining whether outcomes fit theories</p>	<p>B. (1) Enable Program staff to make changes to improve program effectiveness (2) Determining how program is impacting participants (3) Document whether program objectives have been met (4) Facilitate program management</p>
	<p>C. (1) Foster accountability (2) Enable Program staff to make changes to improve program effectiveness (3) Determining how program is impacting participants (4) Facilitate management of programs (5) Document whether program objectives have been met (6) Determining whether outcomes fit theories</p>	<p>C. (1) Determining how program is impacting participants (2) Enable Program staff to make changes to improve program effectiveness (3) (facilitating program management, determining whether program objectives have been met) LAST: Fostering accountability</p>	<p>C. (1) Enable Program staff to make changes to improve program effectiveness (2) Determining impact on participants (3) Determine whether outcomes fit theories (4) Documenting whether program objectives have been met. (2) Foster accountability (3) Facilitate program management.</p>

Table 9—Continued

Categories	Epistemic Styles		
	Metaphoric	Eclectic	Rational
Audiences	A. Stakeholders. (In reality it is the funder. "tyranny of the donor." B. Project staff then stakeholders C. Stakeholders, practitioners, administrators, broader community. In reality it is the funders.	A. Project staff B. Funders, other foundations C. Administrators and funders	A. Greater Public- community awareness B. Key volunteers, Board of Directors, staff. C. Faculty, Administrators, students
Role of the Evaluator	A. 1) change agent, 2) facilitator, 3) interpreter 4) coach, 5) collaborator, educator, consultant, 6) describer, translator 7) investigator, 8) negotiator, mediator, 9) novice B. 1) facilitator, coach, consultant, 2) change agent, educator, patron 3) investigate, interpreter, translator, 4) theorist C. 1) leader, change agent, facilitator, coach, educator, or collaborator Not used by any: Power merchant, connoisseur, expert, critic, judge	A. (1) change agent, coach, investigator, educator, 2) facilitator B. 1) investigator, translator, facilitator, 2) consultant, coach, theorist, interpreter, educator, describer, change agent, collaborator, negotiator, expert, 3) resource C. 1) investigator, change agent, facilitator, collaborator, 2) describer, interpreter, coach 3) mediator, negotiator 4) consultant, expert Not used by any: power merchant, connoisseur, judge, patron, theorist	A. 1) educator, 2) investigator, interpreter, describer, translator, 3) change agent, collaborator, consultant, facilitator, mediator B. 1) educator, 2) theorist, 3) investigator, 4) coach, 5) collaborator, 6) facilitator, 7) interpreter, 8) translator 9) describer, 10) change agent, 11) consultant C. 1) coach, 2) facilitator, 3) educator 4) consultant, 5) change agent, 6) collaborator, 7) theorist, 8) translator, 9) interpreter, 10) investigator

Table 9—Continued

Categories	Epistemic Styles		
	Metaphoric	Eclectic	Rational
Role of Clients	<p>A. Lead the evaluation (theoretically).</p> <p>D. Actively, early, often, thoughtfully</p> <p>E. I would like them to be evaluators too.</p>	<p>A. Learning role</p> <p>B. Community engagement component required</p> <p>C. Little</p>	<p>A. I can't control it.</p> <p>B. Minimal role, would like more</p> <p>C. Yes, throughout process</p>
Role of Stakeholders	<p>A. Yes, theoretically</p> <p>B. Theoretically, yes.</p> <p>C. Definitely</p>	<p>A. It's hard</p> <p>B. Community engagement</p> <p>C. Active role. A source of information. Using them to shape what I'm looking for. Using them to help interpret results.</p>	<p>A. No</p> <p>B. Difficult</p> <p>C. Presently no (stakeholders-students). in prior work with community groups –yes.</p>
Primary Methods of Data Collection	<p>A. Intuitive methods (listen, file away, pull-back, sort, add-in outside documents</p> <p>B. Document analysis (1), interviews (2), focus groups (3), self-developed surveys (4), ethnographic methods (5), standardized instruments (6)</p> <p>C. Observations using ethnographic methods (1), interviews (2), and combination of the rest of the methods (3)</p>	<p>A. Uses them all but would like to do more document analysis such as student portfolios.</p> <p>B. Standardized instruments (1), self-developed surveys, (2), focus groups (3), observations using an instrument (4), observations using ethnographic methods (5), interview (6).</p> <p>C. Focus groups (1), interviews (2), self-developed surveys (3), observations using ethnographic methods (4).</p>	<p>A. Document analysis (1), standardized surveys (2), self-developed surveys(3), focus groups</p> <p>B. Depends on age level of stakeholders. Self-developed surveys, focus groups, interviews</p> <p>C. Standardized surveys, observations using instruments, focus groups , interview, document analysis, observations using ethnographic methods.</p>

Table 9—Continued

Categories	Epistemic Styles		
	Metaphoric	Eclectic	Rational
When is Analysis Planned?	A. Evolves B. Evolves C. Evolves	A. When data collection methods are determined. B. Evolves C. Evolves	A. Both before and after B. After C. Before
Types of Analyses	A. Case studies (1), like lots of information B. Case studies (1), content analysis of qualitative data (2), descriptive-sums, avgs, % (3) C. Case studies (1), content analysis of qualitative data (2), descriptive-sums, avgs, % (3)	A. Descriptive-sums, avgs, %; content analysis of qualitative data, and the study of relationships using cross-tabs and correlations, (1); inferential statistics (2) B. Content analysis of qualitative data (2), case studies (3), descriptive-sums, avgs, % (4), inferential statistics (5) C. Content analysis of qualitative data (1), descriptive-sums, avgs, % (2); studying relationships through cross tabs and correlation (3); case studies- would like to but usually can't because of lack of information, time, or resources.	A. Descriptive-sums, avgs, % (1); content analysis of qualitative data (2), would like to do more studying of relationships with cross-tabs & correlations. B. Content analysis of qualitative data, descriptive-sums, avgs, %, and descriptive case studies. C. Inferential statistics; descriptive-sums, avgs, %; and studying relationships through cross-tabs and correlations (1), content analysis of qualitative data (2) and case studies (3).

Table 9—Continued

Categories	Epistemic Styles		
	Metaphoric	Eclectic	Rational
Communication of Reports	<p>A. Not applicable</p> <p>B. Stories, case studies, videos (using actors). Written report only if necessary.</p> <p>C. Combination of written reports and presentation. Formative evaluation presentations during project.</p>	<p>A. Written reports (1), presentations (2)</p> <p>B. Written reports</p> <p>C. Presentations (1), written reports (2).</p>	<p>A. Written reports (1), press releases (2), presentations (3)</p> <p>B. Written reports (1) occasionally presentations(2)</p> <p>C. Dialogue throughout the duration of the project, then the written report</p>
Sharing of Draft Reports	<p>A. Not applicable</p> <p>B. Yes</p> <p>C. Yes</p>	<p>A. Yes, in various forms</p> <p>B. Yes</p> <p>C. No-do not like to accommodate clients' wishes. Seems to invalidate data.</p>	<p>A. Yes, with anyone. We're a public agency</p> <p>B. Yes, both within and with outside evaluators to get advice</p> <p>C. Yes, feedback is always solicited before final reports</p>
Recommendations	<p>A. NA</p> <p>B. Identification of challenges, Some grounding in what might come next.</p> <p>C. Definitely.</p>	<p>A. It's hard. Try to involve the clients in coming up with their own recommendations.</p> <p>B. Definitely. It should be one of the initial goals.</p> <p>C. Yes</p>	<p>A. If I can</p> <p>B. Yes. What steps could be next. Focus on the logic model</p> <p>C. Yes</p>
Use of Data	<p>A. Yes</p> <p>B. Very much</p> <p>C. Definitely</p>	<p>A. Yes</p> <p>B. Yes</p> <p>C. Yes</p>	<p>A. I hope so</p> <p>B. Yes, yes.</p> <p>C. Yes</p>

Table 9—Continued

Categories	Epistemic Styles		
	Metaphoric	Eclectic	Rational
Ethical Dilemmas	<p>A. Organization bowing to the wishes of the funder</p> <p>B. Reporting weaknesses as well as strengths.</p> <p>C. Being called in to do an evaluation of a program that doesn't exist in any meaningful way. Accountability.</p>	<p>A. Political aspect. Treading carefully. Fine line between truth and half-truth.</p> <p>B. No</p> <p>C. Evaluation for the sake of accountability without caring about meaning. Evaluation for the sake of generating paper better than it is.</p>	<p>A. Guarding of information (defense of territory), confidentiality</p> <p>B. Confidentiality</p> <p>C. Conflict of interests. Evaluators are often paid by the organization being evaluated. Its OK to fail, learning from the failure.</p>
Case Study	<p>A. Would not take the contract. I would decidedly want to be in it in the same time frame as the organizations. I also would want it to be an evaluation to be done by and with the organizations. I would want to act in a very facilitative fashion rather than to do things from an outside evaluator which I'm not interested in doing. I'm committed to the evaluation for learning piece.</p> <p>B. Conference, clustering, development of evaluation questions, interim reports</p> <p>C. Product impact approach, logic models, electronic conferencing, possible cluster evaluation</p>	<p>A. Hold a conference among the 6 projects to discuss logic models and program outcomes. Review proposals and determine probable impact. Make site visits. Have projects report out and document their process.</p> <p>B. Hold a phone conference to talk about criteria and expectations. Make individual phone interviews. Develop profile of projects. Study nature of materials-needs assessment, profile of communities, pair the projects up in clusters. Stakeholder involvement.</p> <p>C. I would be running away from it. I don't have the experience to facilitate this type of evaluation.</p>	<p>A. First gather information about communities-understand their contexts. Make site visits to get a feel for what is happening. Study the plan for implementation. Make analysis of outcomes.</p> <p>B. I don't know. I don't have the type of experience to be able to do this type of work.</p> <p>C. Prefer not to answer.</p>

BIBLIOGRAPHY

- Adelman, C. (1996). Anything goes: Evaluation and relativism. Evaluation, 2, 291-305.
- Affholter, D. P. (1994). Outcome monitoring. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer, Eds. Handbook of practical program evaluation. San Francisco: Jossey-Bass Publishers.
- Alasuutari, P. (1996). Theorizing in qualitative research. Qualitative Inquiry, 2, 371.
- Alkin, M. C. (1990). Debates on evaluation. Newbury Park, CA: Sage Publications.
- Allport, G. W., Vernon, P. E. , & Lindzey, G. (1960). Test manual for the study of values. Boston: Houghton Mifflin Company.
- Altheide, D. L. & Johnson, J. M. (1994). Criteria for assessing interpretive validity in qualitative research. In N. K. Denzin & Y. S. Lincoln (Ed.) Handbook of qualitative research. (pp.485-499). Thousand Oaks, CA: Sage Publications.
- Averch, H. A. (1994). The systematic use of expert judgment. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer, Eds. Handbook of practical program evaluation. San Francisco: Jossey-Bass Publishers.
- Banks, J. A. (1993). The canon debate, knowledge construction, and multicultural education. Educational Researcher, 22, 4-14.
- Barger, R. R. & Duncan, J. K. (1990). Creative endeavor in Ph.D. research: Principles, contexts and conceptions. Journal of Creative Behavior, 24, 59-71.
- Barley, Z. A. & Jenness, M. (1995). Guidelines for the conduct of cluster evaluation. Presentation at AEA Annual Conference, Vancouver, British Columbia, Canada.
- Bickman, L. (1987). The function of program theory. In L. Bickman (Ed.). Using program theory in evaluation. San Francisco, CA: Jossey-Bass.
- Bloom, L. (1998). Under the sign of hope: Feminist methodology and narrative interpretation. Albany, NY: State University of New York.

- Bolman, L. G. & Deal, T. E. (1993). Everyday epistemology in school leadership: Patterns and prospects. In P. Hallinger, K. Leithwood, & J. Murphy, (Eds.). Cognitive perspectives on educational leadership. New York: Teachers College Press.
- BonJour, L. (1985). Structure of empirical knowledge. Cambridge, MA: Blackwell Publishers.
- Burns, J. M. (1978). Leadership. New York: Harper & Row.
- Chen, H. (1990). Theory-driven evaluation. Newbury Park, CA: Sage Publications.
- Coan, R. W. (1979). Psychologists: Personal and theoretical pathways. New York: Irvington.
- Cook, T. D. (1985). Postpositivist critical multiplism. In L. Shotland and M. M. Mark (Eds.). Social science and social policy. Beverly Hills, CA: Sage Publications.
- Creswell, J. W. (1994). Research design: Qualitative and quantitative approaches. Thousand Oaks, CA: Sage Publications.
- Cronbach, L. J. (1982). Designing evaluations of educational and social programs. San Francisco: Jossey-Bass.
- Daly, K. (1997). Re-Placing theory in ethnography: A postmodern view. Qualitative Inquiry, 3, 343-365.
- Dancy, J. & Sosa, E. (Eds.) (1992). A companion to epistemology. Cambridge, MA: Blackwell Publishers.
- Denzin, N. K. (1994). The art and politics of interpretation. In N. K. Denzin & Y. S. Lincoln (Eds.) Handbook of qualitative research (pp. 500-515). Thousand Oaks, Ca: Sage Publications.
- Denzin, N. K. & Lincoln, Y. S. (Eds.). (1994). Handbook of qualitative research. Thousand Oaks, CA: Sage Publications.
- Dewey, J. (1970). Experience and education. New York: The Macmillan Co. (Original work published 1938)
- Eisner, E. (1991). The enlightened eye. New York: Macmillan Publishing Co.
- Eisner, E. W. & Peshkin, A. (1990). Qualitative inquiry in education: The continuing debate. New York: Teachers College Press.

- Everitt, A. (1996). Developing critical evaluation. Evaluation, 2, 173-188.
- Fetterman, D. (1996). Empowerment evaluation: An introduction to theory and practice. In D. M. Fetterman, S. J. Kaftarian, & A. Wandersman, (Eds.). Empowerment evaluation: Knowledge and tools for self-assessment and accountability. Thousand Oaks, CA: Sage Publications.
- Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (Eds.) (1996). Empowerment evaluation: Knowledge and tools for self-assessment and accountability. Thousand Oaks, CA: Sage Publications.
- Fine, M. (1994). Working the hypens: Reinventing self and others in qualitative research. In N. K. Denzin & Y. S. Lincoln, (Eds.). Handbook of qualitative research. (pp. 70-82). Thousand Oaks, CA: Sage Publications.
- Florence, N. (1998). bell hooks' engaged pedagogy: A transgressive education for critical consciousness. Westport, CT: Bergin and Garvey.
- Garrison, J. (1999). Reclaiming the logos, considering the consequences, and removing the context. Educational Theory, 49, 317-337.
- Gergen, K. J. (1985). The social constructionist movement in modern psychology. American Psychologist, 40, 266-275.
- Government Performance and Results Act, (1993).
<http://www.financenet.gov/financenet/fed/cfo/gpra/gpralaw.htm>
- Greene, J. C. (1996). Qualitative evaluation and scientific citizenship. Evaluation, 2, 277-289.
- Greene, J. C. & Caracelli, V. J. (1997). Defining and describing the paradigm issue in mixed-method evaluation. New Directions for Evaluation, 74, 5-17.
- Greenwood, J. D. (1990). Two dogmas of neo-empiricism: The theory-informity of observation and the Quine-Duhem thesis. Philosophy of Science, 57, 553-574.
- Guba, E. & Lincoln, Y. S. (1989a). Effective evaluation. San Francisco: Jossey-Bass
- Guba, E. & Lincoln, Y. S. (1989b). Fourth generation evaluation. Newbury Park, CA: Sage Publications.
- Gubrium, J. F. & Silverman, D. (Eds.) (1989). The politics of field research: Beyond enlightenment. Newbury Park, CA: Sage Publications.

- Haack, S. (1993). Evidence and inquiry. Cambridge, MA: Blackwell Publishers.
- Hamilton, D. (1994). Traditions, preferences, and postures in qualitative research. In N. K. Denzin & Y. S. Lincoln, (Eds.). Handbook of qualitative research. (60-69). Thousand Oaks, CA: Sage Publications.
- Harding, S. (1987). Introduction: Is there a feminist method? In her Feminism and methodology: Social science issues, 1-14. Bloomington: Indiana University Press.
- Heshusius, L. & Ballard, K. (Eds.) (1996). From positivism to interpretivism and beyond. New York: Teachers College Press.
- hooks, b. (1989). Talking back: Thinking feminist, thinking black. Boston: South End.
- House, E. R. (1993). Professional evaluation: Social impact and consequences. Newbury Park: Sage Publications.
- House, E. & Howe, K. (1999). Values in evaluation and social research. Thousand Oaks, CA: Sage Publications
- Huberman, A. M. & Miles, M. B. (1994). Data management and analysis methods. In N. K. Denzin & Y. S. Lincoln (Eds.). Handbook of qualitative research. (pp. 428-444). Thousand Oaks, CA: Sage Publications.
- Johnson, M. (1999, June 10). Senate approves merit grants. Detroit Free Press. (<http://www.freep.com/news/mich/qmerit10.htm>).
- Kasten, K. L. & Ashbaugh, C. R. (1991). The place of values in superintendents work. Journal of Educational Administration, 29, 54-66.
- Keirse, D. & Bates, M. (1978). Please understand me: Character and temperament types. Del Mar, CA: Prometheus Nemesis Book Co.
- Kohn, A. (1999, December 9). Tests that cheat students. The New York Times, pp. 1.
- Kuhn, T. S. (1962). The structure of scientific revolutions. Chicago: University of Chicago Press.
- Lather, P. (1993). Fertile obsession: Validity after poststructuralism. The sociological quarterly, 34, 674-693.
- Lather, P. (1991). Getting smart: Feminist research and pedagogy with/in the postmodern. New York: Routledge.

- LeCompte, M. D. & Goetz, J. P. (1982). Problems of reliability and validity in ethnographic research. Review of Educational Research, 52, 31-60.
- Lewis, C. I. (1952). The given element in empirical knowledge. Philosophical Review, 61, 168-175.
- Lincoln, Y. S. (1990). The making of a constructivist. In E. G. Guba (Ed.) The paradigm dialogue. (pp. 67-87). Newbury Park, CA: Sage Publications.
- Lincoln, Y. S. & Guba, E. G. (1990). Judging the quality of case study reports. Qualitative Studies in Education, 3, 53-59.
- Lincoln, Y. S. & Guba, E. G. (1985). Naturalistic Inquiry. Beverly Hills, CA: Sage Publications.
- Lowry, R. J. (1974). A. H. Maslow: An intellectual portrait. Monterey, CA: Brooks/Cole Publishing Co.
- Madriz, E. I. (1998). Using focus groups with lower socioeconomic status Latina women. Qualitative Inquiry, 4, 114-128.
- Marais, L. (1998). The relationship between organizational culture and the practice of program evaluation in human service organizations. Doctoral dissertation, Western Michigan University, (1998). Dissertation Abstracts International
- Marshall, C. & Rossman, G. B. (1995). Designing qualitative research. Thousand Oaks, CA: Sage Publications.
- Maxwell, J. A. (1992). Understanding and validity in qualitative research. Harvard Educational Review, 62, 279-300.
- McGettigan, T. (1997). Uncorrected insight: Metaphor and transcendence "after truth" in qualitative inquiry. Qualitative Inquiry, 3, 366-383.
- Merriam-Webster's Collegiate Dictionary. 10th Ed. (1997). Springfield, MA: Merriam-Webster, Incorporated.
- Miles, M. B. & Huberman, A. M. (1994). Qualitative data analyses. Thousand Oaks, CA: Sage Publications.
- Miller, D. (Ed.) (1985). Popper selections. Princeton, NJ: Princeton University Press.

- Miller, S. M., Nelson, M. W., & Moore, M. T. (1998). Caught in the paradigm gap: Lived experience and the politics of epistemology. American Educational Research Journal, 35, 377-416.
- Morse, J. (1994). Qualitative research: Fact or fantasy? In J. Morse (Ed.) Critical issues in qualitative research methods. Thousand Oaks, CA: Sage Publications.
- Moser, P. K., Mulder, D. H., & Trout, J. D. (1998). The theory of knowledge. New York: Oxford University Press.
- Neimeyer, G. J.; Prichard, S.; Lyddon, W. J.; & Sherrard, P. A. D. (1989). The role of epistemic style in counseling preference and orientation. Journal of Counseling and Development, 71, 515-523.
- The New London Group, (1996). A pedagogy of multiliteracies: Designing social futures. Harvard Educational Review, 66, 66-92.
- Norusis, M. J. (1998). SPSS for Windows Base System User's Guide Release 6.0. Chicago, IL: SPSS, Inc.
- Patton, M. (1980). Qualitative evaluation. Beverly Hills, CA: Sage.
- Patton, M. (1997). Utilization focused evaluation. Thousand Oaks, CA: Sage Publications.
- Pawson, R. (1996). Three steps to constructivist heaven. Evaluation, 2, 213-219.
- Payne, D. A. & McMorris, R. F. Eds. (1967). Educational and psychological measurement: Contributions to theory and practice. Waltham, MA: Blaisdell Publishing Company.
- Phillips, D. C. (1990). Subjectivity and objectivity: An objective inquiry. New York: Teachers College.
- Piaget, J. (1963). The origins of intelligence in children. New York: W. W. Norton & Company. (Original work published 1952)
- Popkin, R. H. & Stroll, A. (1986). Philosophy made simple. New York: Doubleday.
- Popper, K. (1983). Realism and the aim of science. New York: Routledge. (Original work published 1956)
- Popper, K. R. (1960/1985). Knowledge without authority. In D. Miller (Ed.) (1985). Popper Selections. Princeton, NJ: Princeton University Press

- Popper, K. R. (1970). Normal science and its dangers. In I. Lakatos & A. Musgrave (Eds.), Criticism and the growth of knowledge. Cambridge, MA: Cambridge University Press.
- Popper, K. (1980). Science: Conjectures and refutations. In E. D. Klemke, R. Hollinger, & A. D. Kline (Eds.), Introductory readings in the philosophy of science (pp. 19-34).
- Punch, M. (1994). Politics and ethics in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of qualitative research. (pp. 83-97). Thousand Oaks, CA: Sage Publications.
- Rog, D. J. (1994). Constructing natural "experiments." In J. S. Wholey, H. P. Hatry, & K. E. Newcomer, Eds. Handbook of practical program evaluation. San Francisco: Jossey-Bass Publishers.
- Royce, J. R. (1964a). The encapsulated man. Princeton, NJ: D. Van Nostrand Company.
- Royce, J. R. (1964b). Three ways of knowing and the scientific world view. Methodology and Science, 11, 146-168.
- Royce, J. R. & Mos, L. P. (1975). The Psycho-Epistemological Profile. Edmonton, Alberta: University of Alberta.
- Royce, J. R. , Mos, L. P., and Kearsley, G. P. (1975). Manual: Psycho-Epistemological profile. Edmonton, Alberta: University of Alberta.
- Saloman, G. (1991). Transcending the qualitative-quantitative debate: The analytic and systematic approaches to educational research. Educational Researcher, 20 (6) p.10-18.
- Schon, D. A. (1983). The reflective practitioner: How professionals think in action. New York: Basic Books.
- Scriven, M (1973). The methodology of evaluation. In B. R. Worthen & J. R. Sanders, Educational evaluation: Theory and practice. Belmont, CA: Wadsworth.
- Scrivin, M. (1996). The theory behind practical evaluation. Evaluation, 2, 393-404.
- Shaddish, W. R., Cook, T. D., & Leviton, L. C. (1991). Foundation of program evaluation: Theories of practice. Newbury Park, CA: Sage Publications.

- Spradley, J. P. (1979). The ethnographic interview. New York: Holt, Rinehart, and Winston
- Stanley, L. & Wise, S. (1990). Method, methodology, and epistemology in feminist research processes. In L. Stanley (Ed.) Feminist praxis: Research, theory, and epistemology in feminist sociology. (pp.20-60). London: Routledge.
- Stufflebeam, D. L. (1982). The CIPP model for program evaluation. In G. F. Madaus, M. Scriven, and D. L. Stufflebeam (Eds.) Evaluation models: Viewpoints on educational and human services evaluation. Boston: Kuwer-Nijhoff.
- Swandt, T. A. (1997a). Evaluation as practical hermeneutics. Evaluation, 5, 69-83.
- Swandt, T. A. (1997b). Qualitative inquiry: A dictionary of terms. Thousand Oaks, CA: Sage Publications.
- Taylor, V. & Rupp, L. J. (1991). We make our own history, but not just as we please. In M. M. Fonow & J. A. Cook (Eds.). Beyond methodology: Feminist scholarship as lived research. (pp.119-132). Bloomington: Indiana State University Press.
- Vedung, E. (1997). Public policy and program evaluation. New Brunswick, NJ: Transaction Publishers.
- Wagner, J. (1993). Ignorance in educational research, or how can you not know that? Educational researcher, 22, (5), 15-23.
- Walsh-Sarnecki, P. & Van Moorlehem, T. (1998, Jan 8). Twenty-one schools may face sanctions. Detroit Free Press. (<http://www.freep.com/news/education/accred8.htm>).
- Wandersman (Eds.) Empowerment evaluation: Knowledge and tools for self-assessment and accountability. Thousand Oaks, CA: Sage Publications.
- Weiss, C. H. (1998). Evaluation. Upper Saddle River, NJ: Prentice-Hall, Inc. (Original work published 1972)
- W. K. Kellogg Foundation (1998). W. K. Kellogg Foundation Evaluation Handbook. Battle Creek, MI: W. K. Kellogg Foundation.
- Wong, L. M. (1998). The ethics of rapport: Institutional safeguards, resistance, and betrayal. Qualitative Inquiry, 4, 178-199.
- Worthen, B. R. & Sanders, J. R. (1987). Educational evaluation: Alternative approaches and practical guidelines. White Plains, N.Y: Longman

Yin, R. K., Kaftarian, S. J., & Jacobs, N. F. (1996). Empowerment evaluation at federal and local levels. In D. M. Fetterman, S. J. Kaftarian, and A.

Ysseldyke, S. (1988). Assessment: Assessment in special and remedial education. Fourth edition. Boston: Houghton Mifflin Company.