AN EXAMINATION OF THE DOMINANT AND THE ALTERNATIVE PARADIGMS IN CURRICULUM EVALUATION

By

GLADYS E. PERELES

Bachelor of Science in Secondary Education
Catholic University of Puerto Rico
Ponce, Puerto Rico
1972

Master of Education
Catholic University of Puerto Rico
Ponce, Puerto Rico
1976

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the degree of
DOCTOR OF EDUCATION
May, 1987

AN EXAMINATION OF THE DOMINANT AND THE

ALTERNATIVE PARADIGMS IN

CURRICULUM EVALUATION

Thesis Approved:

J. Randall Koetting
Thesis Adviser

Harold Delaksnicks

Russell Olson

Kenneth St. Clair

Norman N. Durham
Dean of the Graduate College

DEDICATION

To my loving sons, Alexis, Carlos, Pepe, and Pucho, and to my mother, Rosa, and in memory of my father Genare Pereles.

ACKNOWLEDGMENTS

I wish to thank my thesis adviser, Dr. J. Randall Koetting, who suggested that this study be conducted and whose faith provided the most helpful support. Also, I wish to express my appreciation to my major adviser, Dr. Russell L. Dobson, who prodded and encouraged until this study was completed. My appreciation is also extended to the other committee members, Dr. Kenneth St. Clair and Dr. Daniel Selakovich, for their support and understanding.

I am indebted to Eduardo Flores for his assistance in the preparation of the figures, and also for his genuine friendship and invaluable help. Words of indebtedness are also expressed to my friends, Betty Jo McCarty and Jose and Ivette Gonzalez, for their never-failing encouragement, support, and assistance.

A note of thanks is given to Mary Lee Hailey for her patience in typing the first three chapters of this manuscript, and also for her invaluable revisions. I would also like to thank Shirley Motsinger and members of the staff at First Word for editing and preparing the final copy.

Finally, I want to express my eternal debt to my beloved sons, Alexis, Carlos, Pepe, and Pucho Santiago, and

my mother, Rosa Velez, for their understanding and many sacrifices.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Curriculum Evaluation: State of the Field

The field of curriculum evaluation is at present in a state of crisis. This statement is supported by an analysis of crisis provided by the prominent historian and philosopher, Thomas S. Kuhn, in his influential study about paradigm changes in historical development (Kuhn, 1970).

According to Kuhn, a crisis in a field of activity is an indication that the existing paradigm "has ceased to function adequately" (p. 92). This malfunction or "breakdown" of the paradigm becomes apparent by the presence of persistent "anomalies." By anomalies he means discrepancies between the findings produced by the application of the paradigm's tools (i.e., theories, concepts, methods, or instruments) and new sorts of evidence. These discrepancies, as he notes, are characterized by a "stubborn refusal" to be "accommodated" within the existing paradigm (p. 97).

In the curriculum evaluation field a dominant paradigm can be identified. This dominant paradigm comes from the positivist tradition of the natural sciences. It is

referred to here as the natural science paradigm. There is evidence which indicates that curriculum evaluation activities conducted under this dominant paradigm continually produce discrepant data. For example, Guba (1969), in his assessment of the status of the evaluation of curriculum innovations, asserts that

> . . .traditional methods of evaluation have failed educators in their attempts to assess the impact of innovations in operating systems. Indeed, for decades the evidence produced by the application of conventional evaluation procedures has contradicted the experiential evidence of the practitioner. Innovations have persisted in education not because of the supporting evidence of evaluation but despite it (p. 19).

Similarly, Stufflebeam et al. (1971), in their review of the evaluation of instructional programs, conclude that program evaluation consistently "produces findings that are at variance with experience and common observations" (p. 8). Perhaps the most substantial evidence of the presence of anomalous findings in natural-scientific evaluations is the fact that many of those evaluations have been revised. Included in this revision are the Westinghouse/Ohio University Head Start investigation (Berk, 1981; Evans, 1971; Madaus, Airasian and Kellagham, 1980); the evaluations of Sesame Street and Plaza Sesame (Anderson et al., 1977); and several evaluations of Title I curricula funded under the Elementary and Secondary Education Act (ESEA) of 1965 (McLaughlin, 1975).

Although practitioners most committed to the dominant paradigm do acknowledge the presence of anomalies, they do not fault their paradigm. It is maintained that the factor

responsible for such a situation is the poor quality of evaluation practice. This position is best stated by Worthen (1977):

> It should not be surprising that there are poorly planned and executed evaluations; such failings will occur in any field of human endeavor. The problem is one of frequency and importance. So many key evaluations have been disappointing that even some evaluation advocates are beginning to wonder whether evaluation can live up to its high potential.... Unless its practice is improved significantly in the next few years, evaluation will not only fail to meet its potential, but eventually may be discarded as another promising notion that failed to mature to usefulness (pp. 3-4).

Worthen goes on to suggest the establishment of a mechanism of quality control (i.e., standards) for improving evaluation practice. Similar proposals have been advanced by individual evaluators and several professional organizations. In fact, there are already two sets of evaluation standards available. A joint committee appointed by twelve different organizations, including the Association for Supervision and Curriculum Development, developed one set (Stufflebeam and Madaus, 1983). The Evaluation Research Society articulated the other (Rossi, 1982).

Such a position in evaluation is not surprising. As Kuhn explicitly notes, practitioners of a paradigm,

> ...confronted by even severe and prolonged anomalies...do not...treat anomalies as counter-instances, though in the vocabulary of philosophy...that is what they are (p. 77).

This attitude toward anomalies derives, according to Kuhn, from not only practitioners' great confidence in the validity of their paradigm, but more important, from the very

dominance that the paradigm exerts over practitioners. As he puts it, "A paradigm governs, in the first instance, not a subject matter but rather a group of practitioners" (p. 180). Barbour's (1980) brief summary of Kuhn's notion of the power of paradigms provides a good idea of how this dominance is exerted:

> A paradigm...implicitly defines for a given... community the types of question that may legitimately be asked, the types of explanation that are to be sought, and the types of solution that are acceptable. It moulds the scientist's assumptions as to what kinds of entity there are in the world...and the methods of enquiry suitable for studying them (p. 223).

Training, Kuhn suggests, socializes practitioners in the implicit view of the paradigm.

> The study of paradigms...is what mainly prepares the student for membership in the particular ...community with which he will later practice. Because he there joins men who learned the bases of their field from the same concrete models, his subsequent practice will seldom evoke overt disagreement over fundamentals. Men whose research is based on shared paradigms are committed to the same rules and standards for...practice (1970, p. 11).

As suggested above, a breakdown at the level of paradigm is seldom acknowledged by practitioners. But the resulting effect of such a breakdown, which is the crisis, does not depend upon practitioners' recognition of the paradigm's breakdown. This resulting crisis, according to Kuhn, is characterized by four major signs. He identifies these four signs of crisis as (1) expressions of explicit discontent, (2) proliferation of theoretical alternates, (3) discussions and debates over legitimate methods, and (4) emergence of new paradigms. These four signs of

crisis, as the following discussion illustrate, are visible in the field of curriculum evaluation.

Explicit Discontent

The recent literature has already addressed the topic of discontent. For example, Eisner (1985b), in a review of the more general field of curriculum, comment upon the discontent in evaluation in particular. He views this discontent with optimism. He perceives it as a reaction against the "dominance of a scientific epistemo-logy" which has "excluded any other view of the way in which inquiry in education can legitimately be pursued" (p. 17). Eisner points to a group of discontented evaluators and curricularists who are "beginning to look elsewhere" for new modes of rationalities on which to base educational inquiry. Although he agrees with this search, he warns the discontent to try to

> ...avoid the pitfall that so may progressives fell into, both in the 1930s and in the 1960s: namely the tendency toward romantic obscurantism, the infatuation with vague rhetoric that has little intellectual rigor (p. 21).

Jackson (1980) also discusses the topic of discontent. He identifies not only a group of discontented curricularists but also several centers of activity where the discontent is most intense. Among these centers are included the University of Rochester, and Stanford University.

Although Jackson includes himself among these discontents, he disagrees with several arguments advanced by what he calls the "extreme" position within the discontents. He

labels this position as "antiempirical." Such a position, he argues, makes claims about educational matters which they cannot substantiate. He views this extreme position with suspicion:

> ...I suspect that the...extreme critics among our discontents would not be satisfied with the elimination of trivial or poorly conducted empirical studies but, rather, would like to throw out the baby and the bath. In this extreme view, all attempts at quantification are seen as violating the complexity of reality. From this perspective, science--with all its talk about reliability, objectivity, and the rest--is reduced to little more than vain posturing. I strongly disagree with this extreme view and find it troubling (p. 170).

The origins of this discontent can be traced to the national curriculum reform movement of the early 1960s. During this period a number of new curriculum development projects, especially in the areas of science and mathematics, were established. Funds were made available to evaluate these curriculum development efforts. Hyman and Napier's (1975) description of the dominant evaluation approach provides a good idea of how these curriculum development efforts were to be evaluated:

> ...outcomes are hypothesized, an independent variable...the curriculum...is introduced to an experimental population, control groups are identified, and pre and post test data are gathered on both the experimental and control populations. The task of the evaluator is to make sure that the objectives are clear and measurable, to develop the appropriate

> instruments,and to report the results in a systematic manner in terms of which hypotheses have been supported and which should be rejected (p. 58).

It was not so long before developers and implementers of these curriculum efforts were complaining about evaluation results. Most of these new curricula resulted in "no significant difference," results which were in contradiction with what was observed: that these new curricula were making a great difference in the education of children (Guba, 1969; Madaus, Stufflebeam, and Scriven, 1983; Stufflebeam et al., 1971). This complaint and other problems with evaluation were made explicit by Cronbach (1963) in a landmark article on evaluation entitled "Course Improvement Through Evaluation." In his introductory statement, Cronbach expressed:

> ...I am becoming convinced that some techniques and habits of thought of the evaluation specialists are ill-suited to current curriculum studies. To serve these studies, what philosophy and methods of evaluation are required? And particularly, how must we depart from the familiar doctrines and rituals of the testing game? (p. 672).

Cronbach was against the use of norm-referenced test scores as a measure of curriculum effectiveness. According to Cronbach, such scores may produce some "confidence" in a given curriculum, but they will tell "very little about how to produce further improvement" (p. 680). He recommended, instead, the use of item data because he believed that those measures were "more likely...to suggest how to alter the presentation" (p. 60). He was also in disagreement with objective-based evaluations. He suggested the use of

decisions rather than objectives as the advance organizer of the evaluation; i.e., the particular focus for the information to be collected.

Cronbach was one of the first to raise serious questions about the utility and relevance of comparative evaluations. He advised evaluators to turn away from comparisons and to focus instead on the ways in which refinements and improvements could occur while the curriculum was in the process of development. As he put it,

> Evaluation used to improve the course when it is still fluid contributes more to improvement of education than evaluation used to appraise a product already placed on the market (p. 675).

But Cronbach's recommendations went largely unnoticed (Madaus, Stufflebeam, and Scriven, 1983). Subsequent curriculum development efforts, particularly those funded through the ESEA of 1965, received the traditional evaluation treatment. The failure of these evaluations (Guba, 1969) contributed to the spread of discontent in the field of curriculum evaluation. Curriculum evaluators such as Atkin (1968), Eisner (1967), Guba (1969), Scriven (1967), and Stake (1967) were among those evaluators who explicitly expressed their discontent with curriculum evaluation studies in the late 1960s.

During the decade of the 1970s the discontent with curriculum evaluations grew in size and intensity. At present this discontent, borrowing Jackson's (1980) expression, is "sufficiently widespread to be commented upon by any thoughtful observer" (p. 165).

Proliferation of Evaluation Models

In a recent state-of-the-art assessment, Hilliard (1984) describes evaluation as a "hot bed of activity" (p. 124). As he notes, evaluation theories "proliferate almost exponentially" (p. 117). In fact, a review of the theoretical literature (Guba and Lincoln, 1981), reveals that more than forty alternative evaluation models have been already formalized. Such a proliferation of alternates, as Kuhn (1970) notes, signals a field in crisis.

> History...indicates that invention of alternates is just what scientists seldom undertake.... So long as the tools a paradigm supplies continue to prove capable of solving the problem it defines.... The reason is clear...retooling is an extravagance to be reserved for the occasion that demands it. The significance of crisis is the indication they provide that an occasion for retooling has arrived (p. 76).

Several observers of the evaluation scene view the proliferation of evaluation models as creating confusion about the process of evaluation. Gephart (1976, in particular, comments:

> We have reached a point of absurdity! In a recent conversation 27 'different' models of the evaluation process were delineated.... This sorry state of affairs is made even worse by our tendency to refer to the Stake, or Stufflebeam, or Scriven, or Atkin, or Scriven #2, or Provus models... (p. 2).

Gay (1980) expresses:

> A false dichotomy has been fostered in the minds of many to the effect that classroom, or pupil, evaluation and other types of evaluation, such as project evaluation, involve entirely different processes. Further, it is incorrectly believed by a number of people that each type of evaluation requires a different process or model. The

> current proliferation of 'different' evaluation models only serves to reinforce this erroneous belief (p. 14).

Related to this, Gay argues that the evaluation process is the same regardless of what is being evaluated. According to him, what differs in evaluation is the object being evaluated, how the process is applied, and the types of decisions made. Depending upon what is being evaluated, different kinds of data will be collected, different criteria will be applied to the data, and different kinds of decisions will be made.

Classification of Evaluation Models. Several attempts have been made to develop a classification scheme in order to discuss evaluation models. Popham (1975), for example, suggests a four classification-scheme based on emphasis. His four categories are (1) goal-attainment models, (2) judgmental models emphasizing intrinsic criteria, (3) judgmental models emphasizing extrinsic criteria, and (4) decision-facilitation models. According to Popham's scheme, the goal-attainment models predominate in the evaluation field.

Worthen and Sanders (1973) use a three classification-scheme, according to differences in evaluation strategies. Their three categories are (1) judgmental strategies, (2) decision management strategies, and (3) decision-objective strategies. Once they classified contemporary evaluation models within these categories, they proceed to analyze them in terms of twelve selected categories (pp. 210-215).

For Worthen and Sanders, the dominant approach to evaluation, the Tyler model, emphasizes decision-objectives strategies.

Stufflebeam and Webster (1980) classify evaluation approaches according to three categories: (1) politically-oriented, (2) questions-oriented, and (3) values-oriented. Evaluation studies conceived within a politically-oriented perspective are labelled "pseudo-evaluations" (p. 6). They view those types of studies as politically controlled and inspired by public relations. The questions-oriented studies are described as "quasi-evaluations." They noted that these types of studies sometimes are incapable of assessing the worth of educational programs.

> These studies can be called quasi-evaluation studies, because sometimes they happen to provide evidence that can be used to assess the worth of an object; while, in other cases, their focus is too narrow or is only tangential to questions of worth...the main caution is that these types of studies not be equated with evaluation (p. 8).

Two dominant types of evaluation studies, the objective-based and the experimental research, are categorized as quasi-evaluations. In contrast, evaluation studies which they view as effective in assessing the worth of educational programs are those conducted within a value orientation; e.g., decision-oriented and client-centered studies. These studies are considered "true evaluations."

Guba (1978) employs a two-category system to differentiate among evaluation models. His categories are based on the forms of inquiry involved, either conventional or naturalistic. According to Guba, conventional inquiry stems

from a positivist epistemology while naturalistic inquiry derives from phenomenology. Having established the epistemological bases of the two categories, he contrasts evaluation models along thirteen dimensions. These dimensions are: (1) inquiry paradigm, (2) purpose, (3) stance, (4) framework/design, (5) style, (6) reality manifold, (7) value structure, (8) setting, (9) context, (10) conditions, (11) treatment, (12) scope, and (13) methods. Most evaluation models, Guba suggests, fall within the conventional category.

House (1979) proposes a two-classification scheme based on a taxonomy of the most prominent evaluation models (pp. 4-5). He categorizes those models either utilitarian or intuitionist-pluralists. According to House, both categories of models are based on variations in the assumptions of the ideology of liberalism. He compares the utilitarian and intuitionist-pluralist models in terms of their ethical, epistemological, and political assumptions. He views the utilitarian models as dominating the evaluation field.

In a recent work, Guba and Lincoln (1981) suggest that alternative evaluation models can be classified into schools based on what they take as advance organizers. They propose five organizers as categories: (1) objectives, (2) decisions, (3) effects, (4) critical guidespots, and (5) issues and concerns. This latter category, issues and concerns, is the advance organizer of the responsive approach they particularly advocate.

This book offers a new model of evaluation...

> that organizes evaluation activities so as to provide information that illuminates the claims, concerns, and issues raised by stakeholding audiences, that is, audiences involved with or affected by the 'evaluand' (pp. ix-x).

Methodological Debate

What set of methods is most appropriate for evaluation? This question has generated a heated debate between advocates of quantitative methods and proponents of qualitative methods. In general, the two method-types are viewed as incompatible, and a sharp distinction has been made between the two, particularly on the mode of inquiry employed and the basic purpose of such inquiries.

Quantitative inquiry is described as hypothetico-deductive, particularistic, inferential, verificatory, context-free, and outcome-oriented (Reichardt and Cook, 1979; Patton, 1980). This type of inquiry, as Koetting (1984) explains, aims to

> ...develop a 'nomothetic body of knowledge.' This knowledge is best stated in law-like (nomological generalizations) which are seen as truth statements outside of time and specific context... (p. 9).

These nomological generalizations are used, in turn, to predict and control the phenomenon under investigation.

In contrast, qualitative inquiry is viewed as inductive, holistic, descriptive, exploratory, context-bound, and process-oriented (Reichardt and Cook, 1979; Patton, 1980). This type of inquiry is aimed at developing an "ideographic body of knowledge" that exemplifies the individual case (Koetting, 1984, p. 10). As Von Wright (1971)

suggests, this aim is the concern of the human sciences, e.g., history, while the development of nomothetic knowledge is the interest of the natural sciences.

> All these thinkers (Droysen, Dilthey, Simmel, Weber, Windelband, Rickert, Croce, and Collingwood) reject the methodological monism of positivism and refuse to view the pattern set by the exact natural sciences as the sole and supreme ideal for a rational understanding of reality. Many of them emphasize a contrast between those sciences which, like physics or chemistry or physiology, aim at generalizations about reproducible and predictable phenomena, and those which, like history, want to grasp the individual and unique features of their objects. Windelband coined the label 'nomothetic' for sciences which search for laws, and 'ideographic' for the descriptive study of individuality (p. 5).

There is evidence which indicates that a reconciliation between the two method-types has been attempted. For example, Reichardt and Cook (1979) have advanced the thesis that the two method-types are complementary. They provide a definition of the methods which they believe support their complementary thesis. They define quantitative methods as the "techniques of counting, scaling, and abstract reasoning" (p. 22). They contend that these techniques cannot replace qualitative methods since quantitative understanding presupposes qualitative knowing. According to them, the choice of a statistical model to fit the data, the interpretation of findings, and generalizations to other settings all rely on qualitative knowing. They view qualitative methods as the "techniques of personal understanding, common sense, and introspection" (p. 22).

Patton (1980) proposes as a mode of reconciling the two methods a design strategy of "methodological mixes"

(p. 108). This strategy borrows and combines parts such as measurement data, design, and analysis of both methods. He proposes four mixed forms:

1. Experimental design, qualitative measurement, and case analysis.

2. Experimental design, qualitative measurement, and statistical analysis.

3. A holistic-inductive natural design, qualitative measurement, and statistical analysis.

4. A holistic-inductive natural design, quantitative measurement, and statistical analysis.

According to Patton, these mixed forms must be matched to the nature and the needs of the evaluation problem and setting. Guba (1983) criticizes Patton's proposal as relegating the distinctions between quantitative and qualitative methods to "an epistemological oblivion" (p. 22). He argues that the difference between the two method-types is fundamentally epistemological.

LeCompte and Goetz (1982) suggest a supplementary strategy. For example, they propose the introduction of experimental manipulation in ethnographic field work. They argue that ethnographic techniques tend to ignore the threats to internal and external validity. Supplementing ethnographic techniques with experiments, they believe, can enhance the credibility of ethnographic findings.

In a recent attempt to clarify the methodological debate, Smith (1983) criticizes the complementary position. Smith contends that such a position is obscuring the real

issues of the debate rather than clarifying them. In this regard, he expresses:

> How we go about the process of investigation carries with it serious epistemological consequences. These consequences go to the core of educational and social research. Rather than obscure the issues with polemics and name-callings or accept the unfounded assumption that the methods are complementary, we must insure that the problem is the subject of serious and extended debate, not only among philosophers, but even more important, among practicing educational researchers (p. 13).

In contrast, Reichardt and Cook (1979) insist that assuming an incompatibility stance in this controversy is misleading. They state:

> Treating the method-type as incompatible obviously encourages researchers to use one or the other when it may be a combination of the two that is best suited to research needs. It also paralyzes any attempt at reconciling the differences between the opposing sides of the debate over method-types. For these reasons, the conceptualization of the method-types as antagonistic may be leading astray current methodological debate and practice. It is our view that the paradigmatic perspective which promotes this incompatibility between method-types is in error (p. 11).

As a way of response, Guba (1983) advises practitioners to be cautious with what he calls the "siren-song of compromise." He expresses:

> The siren-song of compromise--dare we heed it? I think not.... I do not urge that you stop up your ears--it is important to recognize the song when you hear it. But great care must be taken to avoid being persuaded by it. The sirens are beautiful and sing well; only heroic measures will suffice to withstand them. I wish all of you Godspeed in the confrontation (p. 30).

Emergence of New Paradigms

As previously indicated, curriculum evaluation is dominated by the natural science paradigm. This dominant paradigm has been challenged by three emergent paradigms identified here as the legal, the anthropological, and the aesthetic. A review of literature in this area reveals that each paradigm has gained a constituency of advocates who claim it is their preferred paradigm the "paradigm of choice."

The Legal Paradigm. During the late 1960s, Egon G. Guba suggested that evaluators might look for assistance in the field of law (cf Worthen and Rogers, 1980). In the early 1970s the first example of a legal evaluation in education was conducted. Further interest in the legal approach to evaluation was stimulated by the conceptual work of Wolf (1974; 1975), Owens (1973), and Levine (1974).

According to Wolf (1974), the legal paradigm can overcome what he sees as three notable shortcomings in current evaluation practice: the quantification and oversimplification of data, the neglect of the human intelligence as a powerful collector and interpreter of data, and the lack of credible procedures. He also contends that the emergence of three developments in evaluation, i.e., the political function of evaluation, evaluative advocacy, and the broadening of acceptable evaluation evidence, makes additional case for considering the legal paradigm as the alternative in evaluation.

In a later work, Wolf (1975) proposes a judicial evaluation model which adopts certain concepts from jury trials and adversary proceedings. His model relies on the law's acceptance of human testimony to clarify and judge complex events. This model, as he explains, is aimed at "producing broad program understanding, exploring the complexity of educational issues, and keeping at least two sides of the truth alive" (pp. 185-186). Specifically, it provides for the "structured consideration of alternative arguments and inferences" in order to "keep the evaluation both intellectually honest and fair" (p. 185).

Owens (1973) also utilizes the technique of adversary proceedings. This technique, according to Owens, is based on the litigation principle of the fight theory, which states that "the facts in a case can be ascertained if each side strives as hard as it can, in a keenly partisan spirit, to bring to the court's attention the evidence favorable to that side" (p. 226). He suggests at least five uses of the adversary model in curriculum evaluation: (1) to explore the implicit values of the curriculum, (2) to estimate the congruence between the curriculum and the school system, (3) to select curriculum materials, (4) to judge the merits of the completed curriculum, and (5) to provide information to decision makers.

A strong advocate of the adversary approach, Levine (1974) advances the thesis that the process of conducting an adversary proceeding in our legal system resembles the process of convincing a group of scientists of the validity

of a study. Such a parallel, he argues, provides evidence of the appropriateness of the adversary technique in evaluation. He explains the parallel in this way:

> ...the scientific enterprise as a whole follows an adversary model.... By an adversary model, I mean that we are dealing with a situation in which there are claims and counterclaims and arguments and counterarguments, each side advanced by an advocate who attempts to make the best possible case for his position.
>
> The scientific community, in form of...a referee...acts as a judge does in a preliminary hearing, deciding whether there is sufficient case made in a particular study to take it to the trial before the scientific community. If published, the particular position asserted in the paper is subject to cross-examination and further probing...critiques of experiments, or an area of study, may all be viewed as attacks on a particular position by advocates of another position (p. 669).

Levine's insistence of the appropriateness of the adversary technique in evaluation has gained opposition. Worthen and Rogers (1980), particularly, do not view this technique as the best pattern in legal evaluations. They make this statement:

> The legal paradigm has intriguing possibilities for some evaluation situations.... We are not inclined, however, to view the legal paradigm as...the best pattern. In fact, one of our greatest concerns is that evaluators will seize on some of the more trivial features of the courtroom and fail to isolate and extract those adversarial aspects that might be most pertinent in education (p. 537).

They also raise the suspicion that adversarial evaluations might "result in a seductive slide into what might be termed an 'indictment mentality,' which can do a disservice both to evaluation efforts and the program being evaluated" (p. 537).

The Anthropological Paradigm. The roots of this paradigm are traced to the Malinowski tradition in anthropological/ethnographic field studies, and the Weberian tradition in phenomenology (Guba, 1978). The literature on the anthropological paradigm, as applied to evaluation, began to appear during the late 1960s. Since the late 1970s, proponents of this paradigm have initiated an intensive campaign to convince the evaluation community of the appropriateness of their paradigm.

The most fervent advocates of the anthropological paradigm maintain that this paradigm is not only in competition but also in opposition with the dominant natural science paradigm (Guba, 1978; Guba and Lincoln, 1981; Parlett and Hamilton, 1977; Patton, 1975; 1978; 1980). Patton (1978) establishes the contrast of the two paradigms in this way:

> Evaluation...is dominated by the largely unquestioned, natural science paradigm of hypothetico-deductive methodology. This dominant paradigm assumes quantitative measurement, experimental design, and multivariate, parametric statistical analysis to be the epitome of 'good' science.
>
> By way of contrast, the alternative to the dominant hypothetico-deductive paradigm is derived from the tradition of anthropological field studies. Using the techniques of in-depth, openended interviewing and personal observation, the alternative paradigm relies on qualitative data, holistic analysis, and detailed description derived from close contact with the targets of study (pp. 203-204).

This type of methodology is termed by Guba (1978) as "naturalistic inquiry;" while he describes the methodology of the dominant natural science paradigm as "conventional

inquiry." He points out six major advantages of naturalistic inquiry: (1) deals with emergent questions of interest, (2) provides an acceptable basis for studying process, (3) provides a better description of treatment-situation interactions, (4) avoids implicit shaping of possible outcomes, (5) recognizes pluralistic values, and (6) optimizes generalizabilty.

Parlett and Hamilton (1977), in a description of their proposed illuminative evaluation approach, suggest the purpose of the anthropological paradigm:

> The model described here, illuminative evaluation, takes account of the wider context in which educational programs function. Its primary purpose is with description and interpretation rather than measurement and prediction. It stands unambiguously within the alternative anthropological paradigm (pp. 9-10).

According to Parlett and Hamilton, the task of illuminative evaluation is to provide a comprehensive understanding of the complex realities surrounding the curriculum. A similar purpose is guiding Stake's (1983) responsive approach. In fact, both approaches converge in many respects. This is clearly affirmed by Stake when he himself declares that Parlett and Hamilton's writings "for the most part are harmonious with mine" (p. 291). Guba (1978), however, has shown greater preference for the responsive approach. According to Guba, this approach "seems very naturally wedded to the methodology of naturalistic inquiry" of both ethnographers and phenomenologists (p. 35).

Several critics point out that there is no need for a competition between the natural science and the anthropological paradigms. It is argued that these paradigms can be used either in conjunction or alone so as to best "fit" the demands of the evaluation problem at hand. Reichardt and Cook (1979), in particular, counsel evaluators to

> ...feel free to change their paradigmatic stance as the need arises. These is no reason to subscribe to one mix of attributes at all times. Rather, in moving from one program to the next or from one study to the next...the paradigmatic stance that is most appropriate...is likely to change. Thus, a researcher's paradigmatic viewpoint should be flexible and adaptive (p. 19).

Strong advocates of the dominant natural science paradigm, however, have raised suspicion that the use of the anthropological paradigm in evaluation might result in promoting a type of inquiry which can be termed "journalistic reporting," "anecdotal," or "impressionistic story-telling."

The Aesthetic Paradigm. This paradigm is supported by a group of curriculum critics. Its internal antecedents can be found in the "reconceptual movement" in the more general field of curriculum. This movement attempts to reconceptualize the curriculum field around human consciousness, personal meaning, and collective action (Giroux, 1981; Pinar, 1978). Huebner (1966), following a critique of the dominance of technical values in curriculum theorizing and practice, suggested the need for viewing and assessing the curriculum aesthetically. Building upon Huebner's suggestion, Mann (1969) formulated the concept of "curriculum criticism," and articulated the task of the

curriculum critic, i.e., to disclose meaning inherent in both the curriculum design and in the choices made by its designers.

Since the early 1970s, Elliot Eisner has become the leading advocate of the aesthetic paradigm in curriculum evaluation. In a recent work, Eisner (1985b), after analyzing the controlling effects of scientific evaluation procedures upon educational practice, proposes a critical evaluation approach called "educational criticism." This approach to evaluation "takes its lead from the work that critics have done in literature, theater, film, music, and the visual arts" (p. 216). In postulating the reason for choosing specifically the fine arts as the basis for his critical approach, Eisner declares:

> The arts are not a second-class substitute for expression, they are one of the major means people throughout history have used both to conceptualize and express what has been inexpressible in discursive terms (p. 226).

He further claims that it is through the arts that

> ...we have the opportunity to participate vicariously in the lives of others, to acquire an empathetic understanding of situations, and therefore to know them in ways that only the arts can reveal (p. 227).

According to Eisner, educational criticism is composed of three major aspects or dimensions: descriptive, interpretive, and evaluative. The descriptive dimension is essentially an attempt to portray or render in artistic language the qualities and meanings of classroom life. The interpretive dimension represents an effort to understand the meaning of classroom events. The evaluative dimension

is concerned with making value judgments of the educational import of what has been described and interpreted. In general, educational criticism is aimed at illuminating the complexities of classroom life. Perhaps Eisner (1985b) explains it best when he states:

> The major aim of such activity is not primarily to discover laws but rather to illuminate, to provide those concerned with education with the kind of understanding that will enhance their own teaching or professional deliberations.... It aims not at the reduction of complexities but at their illumination in order that the factors and qualities that make situations unique as well as general can be understood (p. 380).

The educational criticism approach, as Guba and Lincoln (1981) point out, has the "honor to be the first to break cleanly" with the dominant natural science paradigm (p. 20). But this new perspective in curriculum evaluation, like the legal and anthropological paradigms, has also raised suspicions. In particular, Stufflebeam and Webster (1980) suspect that a critical stance in evaluation might leave ample room for "subjectivity, bias, and corruption" (p. 17). But, as Kuhn (1970) indicates, to look at new candidates for paradigm with suspicion is very typical. In this particular, he comments:

> At the start a new candidate for paradigm may have few supporters, and on occasions the supporters' motives may be suspect. Nevertheless, if they are competent, they will improve it, explore its possibilities, and show what it would be like to belong to the community guided by it. And as that goes on, if the paradigm is one destined to win its fight, the number and strength of the persuasive arguments in its favor will increase (p. 159).

Purpose of the Study

Kuhn, after examining the resolution of several cases of crises, concludes that crises are terminated with a simultaneous decision of rejecting the old paradigm and accepting a new one which revolutionizes the field. But he explicitly notes that the act of judgment leading to that decision is preceded by a period of "extraordinary science." By extraordinary science Kuhn means research directed at examining competing paradigms.

Having established (using Kuhn's analysis) that curriculum evaluation is in a state of crisis, and having identified several competing paradigms (the dominant and the emergents) in this field, the next step is to engage in "extraordinary science" research.

One of the basic purposes of this study is to examine select philosophical, conceptual, and methodological aspects of the dominant natural science paradigm . At the level of philosophy, four assumptions will be considered: (1) the nature of reality, (2) the nature of subject-object relationships, (3) the nature of truth statements, and (4) the nature of value judgments. At the conceptual level, this study will seek answers to two fundamental questions: (1) How does this paradigm conceptualize the curriculum being evaluated: and (2) What model(s) does it provide for evaluating the curriculum? The methodological examination will be concerned with the fundamental techniques of data

collection, design, and analysis that this dominant paradigm provides for curriculum evaluation

As a second purpose, this study will examine one of the emergent alternatives--the aesthetic paradigm. It is hoped that this examination will provide the stimulus for researching the other candidates for paradigm, i.e., the legal and the anthropological.

The paradigm concept has received considerable attention in the literature since the first publication of Kuhn's study in 1962. This concept is used to refer to a particular "Weltanschauung." This term comes from the German language and is translated to English as "world view." World view, in turn, is defined as a "comprehensive conception or apprehension of the world" (Webster's New Collegiate Dictionary, 1979, p. 1321). Several analyses of the structure of paradigms reinforce Kuhn's basic notion that a paradigm contains not only a philosophy but also a linkage to certain theories and methods. This study of paradigms incorporates that notion.

Organization of the Study

This study is organized into five chapters. Chapter I has provided its introductory part. This introduction contains an assessment (using Kuhn's analysis of crisis) of the present state of the field of curriculum evaluation and the purpose of the study. The remaining chapters are organized as follows:

Chapter II includes a description of the distinctive features of the dominant natural science paradigm, an examination of the concept of curriculum and the evaluation process as developed in the context of this paradigm, and a discussion of the most prominent evaluation techniques of data gathering, design, and analysis used by practitioners of this paradigm.

Chapter III provides a discussion of the philosophical world view of the dominant natural science paradigm. This discussion focuses on the four sets of assumptions previously indicated, i.e., the nature of reality, the nature of subject-object relationships, the nature of truth statements, and the nature of value judgments.

Chapter IV contains the examination of the aesthetic paradigm. This examination is intended to bring to focus the paradigm's new ideas concerning the curriculum, the evaluation process, and the methods of evaluation. It is also concerned with identifying the paradigm's philosophical orientation.

Chapter V contains a brief discussion of some of the implications of accepting the aesthetic paradigm in curriculum evaluation, a statement of the conclusions of this study and recommendations for further studies.

CHAPTER II

THE FRAMEWORK OF THE DOMINANT PARADIGM

Distinctive Features

As pointed out in the first chapter, the field of curriculum evaluation is dominated by the natural science paradigm. The notion of paradigm was defined as a world view, and it was also suggested that the dominant world view comes from the positivist tradition. In general, the positivist view (discussed more fully in Chapter III) conceives the world as an objective reality which is known to us through sense-experience and verified by natural science methods. This world view considers theology and metaphysics as primitive ways of viewing the world. The term "natural science" is used here to refer to "any of the sciences (as physics, chemistry, or biology) that deal with matter, energy, and their interrelations and transformations" (Webster's New Collegiate Dictionary, 1979, p. 759).

The application of the dominant natural science paradigm is distinguished by two major features: the use of the scientific method, and the quantification of observations.

The Scientific Method

Originally, the scientific method was based on Aristotelian deductive logic. This type of logic utilizes the categorical syllogism as its basic model or reasoning. Syllogistic reasoning establishes a logical connection between three propositions; that is, a major premise, a minor premise, and a conclusion. A major premise is a self-evident truth or a general axiom. A minor premise is a particular case related to the major premise. A conclusion is the logical consequence of the established relationship of the two premises. With the dominance of positivism, Aristotelian logic has been replaced by the logical method of analysis (symbolic logic) developed by the English mathematicians/philosophers Bertrand Russell (1872-1970) and Alfred North Whitehead (1861-1947).

Symbolic logic, like Aristotelian logic, is a deductive method of analysis; that is, it moves from the general to the specific. Unlike Aristotelian logic, it incorporates hypothetical reasoning. In addition, symbolic logic submits the propositions to a "process of reductionism" on which they are operationalized to their immediately observable content (Culbertson, 1981, p. 33). According to the positivists, symbolic logic is superior to Aristotelian logic. This position is best expressed by Kraft (1983), when he states:

> The new logic, symbolic logic, is far superior to traditional logic, both in terms of content and in terms of form. It not only contains a larger

> number of essential disciplines, but even the old disciplines are treated with greater rigor and more systematically. And with the help of symbolism a form of representation has been found which enables mathematically precise formulations of concepts and propositions and rules governing the latter's combination. Thus, it becomes possible to operate in a purely formal manner... Equivocations are avoided, unnoticed assumptions are revealed, rigorous deduction is guaranteed (p. 18).

The application of the scientific method involves three fundamental steps: deduction, verification, and generalization. Deduction, as noted above, is the process of logical derivation of propositions from generalizations. These logical propositions are then formulated in terms of hypothetical or observational statements. Verification consists of designing and performing laboratory experiments to confirm (or disconfirm) those hypotheses. As Guba notes (1978), the purpose of this step is "to rid the field of certain variables not of interest but having possible effects which would produce erroneous results" (p. 12). Guba also points out that in evaluation, since laboratory controls are more difficult to handle, practitioners of the scientific method often find recourse in the strategy of randomization. He states:

> In the physical sciences it is always possible to effect physical control over such confounding variables, but control is far more difficult to manage in dealing with human behavior. Here the investigator must resort to a randomization strategy which, while not eliminating the confounding variables, at least permits him to estimate their size through statistical processes (p. 12).

The final step, generalization, involves the act of inferring from experimental data law-like explanations that can

"provide dependable bases" for predicting and controlling the phenomenon under investigation (Guba and Lincoln, 1983, p. 319). Generalization rests on the assumption of the "lawfulness of nature," which states that "the closer two events are in time, space, and measured value on any or all dimensions, the more they tend to follow the same laws" (Campbell and Stanley, 1963, p. 17).

The search for generalizations is said to be the central interest of the scientific method in particular and the dominant paradigm in general.

> The dominant paradigm is directed at producing generalizations. The assumption that this is the goal of Science is so deeply ingrained that it is virtually true by definition. I have never seen this assumption questioned in the literature of Scientific Methodology. Science is the search for generalizations (Patton, 1975, p. 35).

In fact, Eisner (1983) traces the present dominance of the natural science paradigm to the strong aspiration of earlier educational scientists, especially behavioral psychologists, to discover the "laws of learning" that would predict and control the outcomes of the school curriculum. Such an aspiration, Eisner notes, was/is reflected on the emphasis on efficiency and effectiveness in the schools. He expresses:

> Laws that would do for educational practitioners what the work of Einstein, Maxwell, and Bohr have done for physicists were the object of the educational scientist's dream. This yearning for prediction through control was, of course, reflected in the desire to make schools more efficient and, presumably, more effective.... This aspiration to discover the laws of learning was allied with the scientific management movement in education... It reflected then, as it does today, the need to

> discover the principles and practices that would give us efficient and effective schools (pp. 335-336).

Further, he asserts that this search for the laws of learning has produced three significant consequences: the objectification of the educational practice, the oversimplification of the educational process, and the de-emotionalization of the language for talking about children, teaching, and educational priorities.

Several evaluation theorists argue that in evaluation it is impossible to generalize in the scientific sense. For example, Cronbach (1975), after examining 20 years of evaluative research, concludes that generalizations are like radioactive substances; that is, they decay and have half-lives.

> Generalizations decay. At one time a conclusion describes the existing situation well, at a later time it accounts for rather little variance, and ultimately is valid only as history (p. 122).

Cronbach counsels practitioners to place emphasis on describing and interpreting particular situations rather than viewing generalization as the aim of inquiry. He expresses:

> Instead of making generalization the ruling consideration...I suggest that we reverse our priorities. An observer collecting data in a particular situation is in a position to appraise a practice or proposition in that setting, observing effects in context.... As he goes from situation to situation, his first task is to describe and interpret the effect anew in each locale, perhaps taking into account factors unique to that locale or series of events.... When we give proper weight to local conditions, any generalization is a working hypothesis, not a conclusion (pp. 125-125).

Similar advice is given by Guba (1978) to the naturalistic evaluator:

> Often naturalistic inquiry can establish at least the 'limiting cases' relevant to a given situation. But in the spirit of naturalistic inquiry he should regard each possible generalization only as a working hypothesis to be tested again in the next encounter and again in the encounter after that. For the N/I evaluator premature closure is a cardinal sin, and tolerance of ambiguity is a virtue (p. 70).

Patton (1980) proposes the "inductive" approach for "grounding" generalizations on the data. He distinguishes this approach as follows:

> ...an inductive approach to evaluation research means that an understanding of program activities and outcomes emerge from experience with the program. Theories about what is happening in a program are grounded in this program experience, rather than imposed on the program a priori based on hypothetico-deductive constructions (p. 41).

But, for the dominant natural science paradigm, the scientific method is the superior mode of arriving at generalizations. This is clearly evident in Durant's (1961) synthesis of Comte's positivist Law of Three Stages:

> In each field of thought the historian of ideas could observe a Law of Three Stages: at first the subject was conceived in the theological fashion, and all problems were explained by the will of some deity--as when the stars were gods, or the chariots of gods; later, the same subject reached the metaphysical stage, and was explained by metaphysical abstractions--as when the stars moved in circles because circles were the most perfect figure; finally the subject was reduced to positive science by precise observation, hypothesis, and experiment, and its phenomena were explained through the regularities of natural cause and effect (p. 352).

Quantification

The dominant paradigm demands that any observation concerning the "qualities" of the object of investigation be expressed in quantitative terms. According to Gay (1980), quantification provides at least four advantages: It (1) permits a more precise and objective description concerning qualities, (2) reduces ambiguity considerably, (3) supplies much more information than the verbal statement, and (4) facilitates comparisons. For quantifying observations, the paradigm relies upon the use of measuring instruments which provide a "standardized framework" on which observations are "fit into" via the assignment of "numerical values" (cf. Patton, 1980, p. 22). In curriculum evaluation, the "test" is the preferred measuring instrument.

In methodological discussions, opponents of quantification argue that in the translation of observations to numbers valuable information is not only destroyed but is altered to fit into the standardized categories. Regarding the latter, Eisner (1985b) observes:

> The quantitative inquirer is obliged to transform the qualities perceived into quantitative terms... In this translation the information is altered; differences between qualities are placed on a common scale to make them comparable. This process requires the use of a coding system--number--that is not structurally analogous to the forms that were initially perceived. The number symbol is a representational rather than a presentational symbol (p. 238).

Eisner goes on to identify poetry and literature as two of the modes that employ "structurally analogous" forms; that

is, presentational symbols. It is also argued that when the object of investigation is the human subject there are certain qualities, e.g., motivation or interest, which are not amenable to quantification. Quantitative thinkers maintain that theoretically all qualities can be quantified. The following argument is frequently used to support their position: "If something exists, it exists in quantities, or amounts; if it exists in quantities, it can be measured" (Gay, 1980, p 9).

Patton (1975) observes that nonquantitative approaches to evaluation frequently stimulate charges of "subjectivity." Nonquantitative evaluators are exhorted to eschew subjectivity and make sure that their evaluations are objective. Objective evaluations are considered to be the only kind of studies worthy of federal support. Subjective evaluations, as he notes, are regarded as unscientific.

> To be subjective means to be biased, unreliable, and non-rational. Subjective data imply opinion rather than fact, intuition rather than logic, impression rather than confirmation (p. 21).

Patton maintains that subjectivity properly construed can become a positive rather than a pejorative term in evaluation. Subjectivity allows the evaluator to get close and involved with the data, thereby developing categories from the data itself. He asserts that to identify objectivity as a virtue of quantification is an ideological statement which functions to preserve and protect its dominance in evaluation. Dominance which needs to be maintained, as

Patton suggests, since quantitative methodologists rank at the top of the methodological status hierarchy. In this particular, he states:

> The art and science of quantification constitutes the very core of the dominant paradigm. To turn words into numbers, historical trends into prediction equations, and the behavior of people into probability tables and standardized regressions--those are the greatest miracles in Science, and to the performers of these miracles go the greatest of all scientific rewards: recognition and high status (p. 12).

Patton then argues that the dominance of quantification works in practice to severely limit the kinds of questions to be asked and the types of problems to be studied.

Preferred Conceptions of the Curriculum

The evaluation of the curriculum usually involves internal and external assessments of effectiveness. Internally, evaluation assesses the degree to which the curriculum has achieved its intended outcomes. Externally, evaluation assesses the effects of the curriculum on student achievement. It seems evident that the dominant framework prefers a specific conception of the curriculum for each type of assessment. For example, internal evaluation tends to incorporate a product conception of the curriculum, while external evaluation conceives the curriculum as a treatment.

Curriculum as a Product

The product conception of the curriculum can be noted, for instance, in the tendency of practitioners to use the terms "curriculum" and "product" as synonymous when defin-

ing evaluation-related concepts. Unruh and Unruh's (1984) definition of formative evaluation illustrates this tendency:

> ...formative evaluation refers to the collection of appropriate evidence during the construction and trial of a new curriculum so that revisions of the curriculum can be based on this evidence. formative evaluation is almost exclusively aimed at improving the...product during its developmental phases (p. 273).

Similarly, Zais (1976), in defining summative evaluation, refers to the curriculum as a product:

> Summative evaluation, as its name implies, is conducted in order to obtain a comprehensive assessment of the quality of a completed curriculum. Thus, summative evaluation ordinarily takes place at the completion of the curriculum development process and provides a terminal judgment on the completed product in overall, general terms (p. 381).

This product view of the curriculum stems from the scientific-management movement of the early 1900s that sought to install industrial practices in the schools. A leading figure of the movement, Franklin Bobbitt, delineated a production model for curriculum-thinking based on the notion of efficiency as espoused by the founder of scientific management, the industrial psychologist Frederick W. Taylor. In Bobbitt's (1935) model, the curriculum consisted of four basic ingredients: (1) a statement of objectives, (2) a sequence of experiences shown by analysis to be reasonably uniform in value in achieving the objectives, (3) subject matter found to be reasonably uniform as the best means of engaging in the experiences, and (4) statements of immediate outcomes of achievements to be

derived from the experiences (p. 163). This product was to be refined by evaluation before placing it in the market. Huebner (1975) puts it this way:

> Evaluation...may be considered a type of quality control. The end product is scrutinized to see if it can go to the market with the stamp of approval; or if not yet at the end of the production line, the inadequate products-in-process are shunted aside to be reworked by remedial efforts until they can return to the normal production line (p. 223).

This product view of curriculum was later recaptured by Ralph W. Tyler in his syllabus Basic Principles of Curriculum and Instruction (1949), widely known as the Tyler rationale. This rationale, as Kliebard (1978) observes, has been "raised almost to the status of revealed doctrine" (p. 256).

Tyler's rationale revolves around four basic questions that he feels need to be answered in developing a curriculum. These questions are as follows:

1. What educational purposes should the school seek to attain?

2. What educational experiences can be provided that are likely to attain these purposes?

3. How can these educational experiences be effectively organized?

4. How can we determine whether these purposes are being attained?

These questions may be reformulated into the familiar four-step process by which a curriculum is developed: stating

objectives, selecting experiences, organizing experiences, and evaluating.

According to Tyler (1949), the most crucial step in developing a curriculum is to answer the first question since all the others proceed from this step. As he puts it, "If we are to study an educational program systematically and intelligently we must first be sure as to the educational objectives aimed at" (p. 3). He contends that the most useful way to state educational objectives is by expressing them in behavioral terms. These objectives also become the criteria by which the means to accomplish the purposes are selected, e.g., teaching materials, content, instructional procedures, and evaluation instruments.

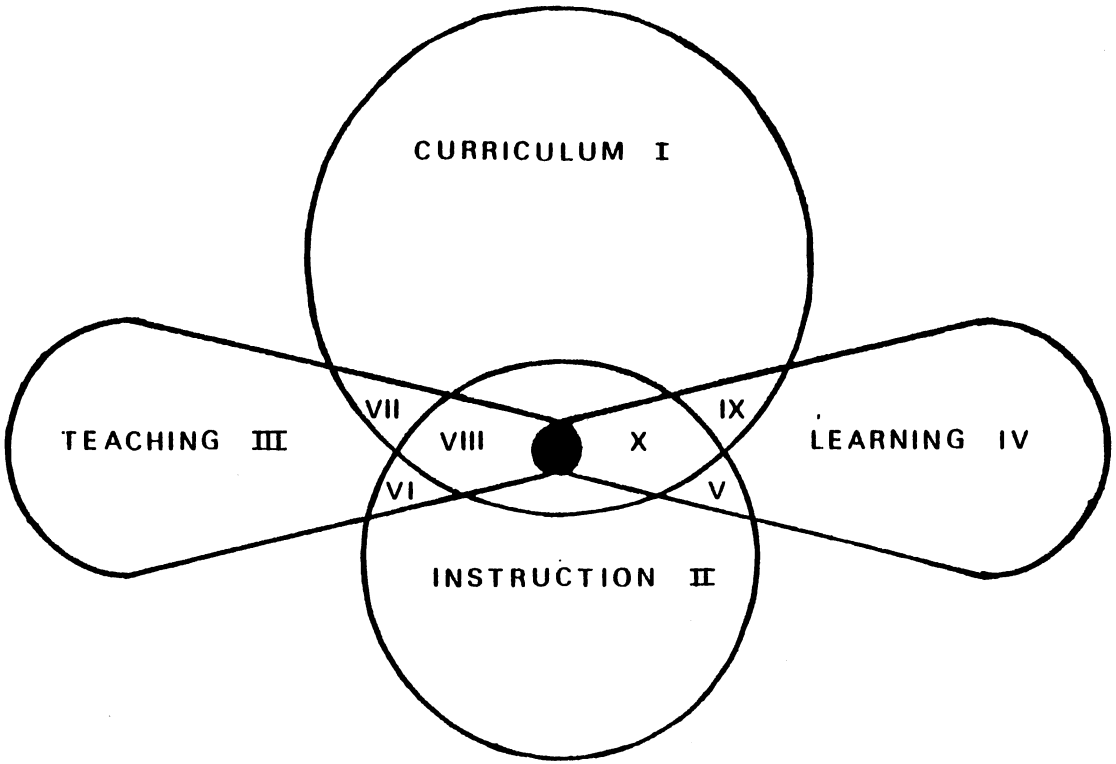
Curriculum Controversy. The inclusion of educational or learning "experiences" in the curriculum has generated much discussion in the curriculum field. It is argued that the statement of specific objectives is what the curriculum is all about. Anything else, e.g., learning experiences, belongs to the realm of instruction. In a widely debated essay, Johnson (1967) points out that "there is...no experience until an interaction between the individual and his environment actually occurs. Clearly, such interaction characterizes instruction, not curriculum" (p. 44). He argues that if curriculum must play any role in instruction, it must be viewed as "anticipatory" not "reportorial" (p. 44). In this regard, he states:

> Curriculum prescribes (or at least anticipates) the results of instruction. It does not prescribe the means...to be used in achieving the

> results. In specifying outcomes to be sought, curriculum is concerned with ends, but at the level of attainable learning products, not at the more remote level at which these ends are justified (p. 44)

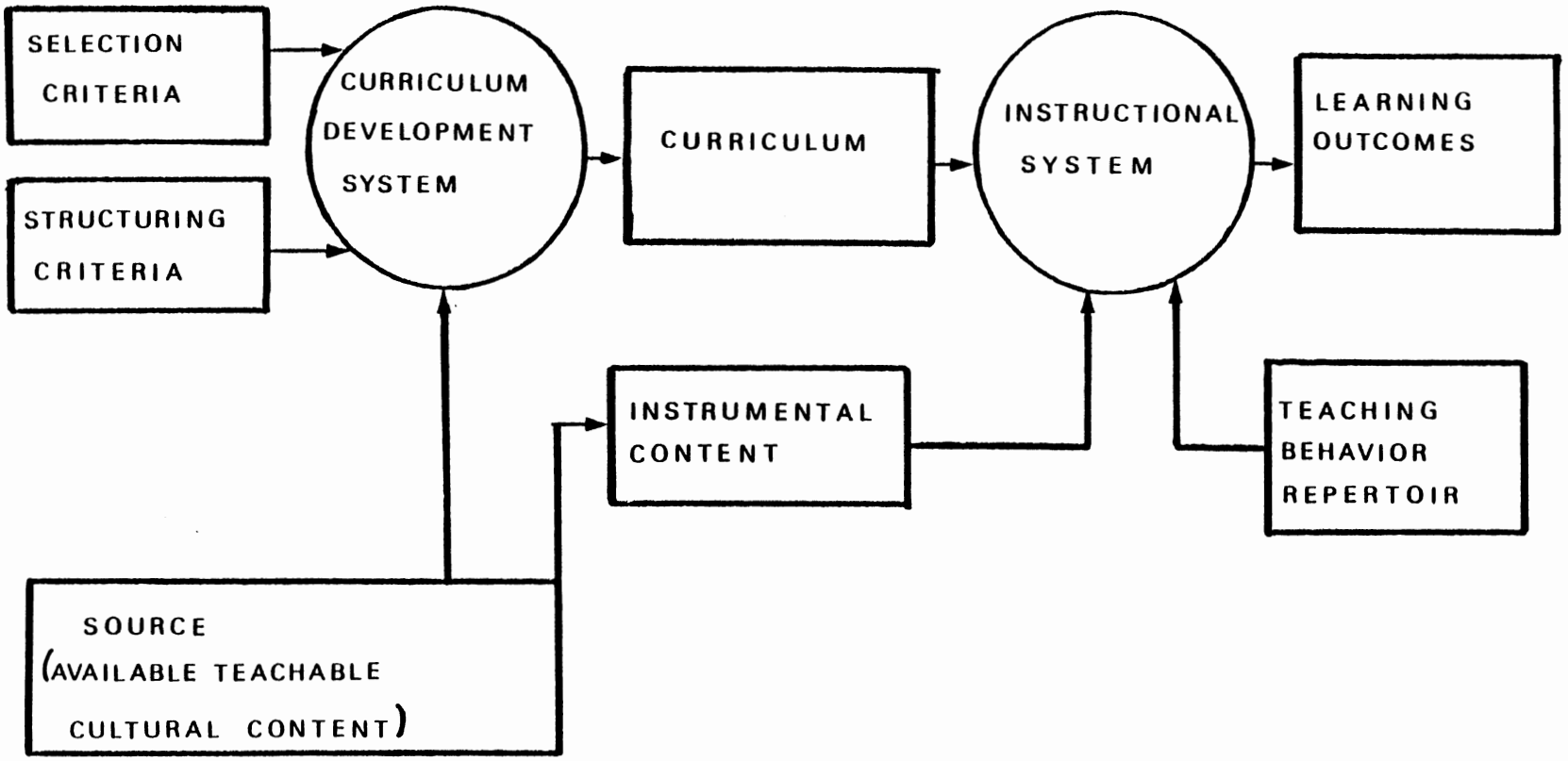
He goes on to define the curriculum as "a structured series of intended learning outcomes" (p. 44). By a "structured series" he means that the curriculum indicates organizational relationships among the "intended outcomes." These intended outcomes, he argues, need to be selected from the available, teachable cultural content (p. 45).

MacDonald (1965), however, contends that the curriculum is not an isolated element in the school system. He views the curriculum as interacting not only with instruction but also with teaching and learning. He presents a systems model of curriculum showing this interaction (Figure 1). In his model, teaching is defined as a personality system (the teacher) acting in a professional role. Learning is another personality system (the student) who is perceived by the teacher as performing task-oriented activities. Instruction is the combination of these two personality systems, which also may be thought of as the "action context" within which formal teaching and learning take place. This combination, according to MacDonald, eventuates in a social system. The curriculum is also a social system. This social system consists of those individuals whose actions eventuate in "a curriculum." He then defines "a curriculum" as essentially a plan for action; that is, a plan which guides instruction (1965, p. 6).



Source: J. B. MacDonald. Educational models for instruction. In J. B. MacDonald and R. R. Leeper (Eds.), Theories of Instruction, 1965.

Figure 1. MacDonald's Model of Curriculum



Source: M. Johnson. Definitions and models in curriculum theory, Educational Theory, 1967.

Figure 2. Johnson's System Model of Curriculum

In response to MacDonald's contention, Johnson (1967) argues that the curriculum is not a system itself but the output of one system (the curriculum development system) and an input into another system (the instructional system). He also supports his argument with a system model of curriculum (Figure 2). Johnson also claims that curriculum evaluation is too often confounded with instructional evaluation. According to him, curriculum evaluation involves validation of the selection and the structure of the curriculum. The curriculum serves as the criterion for instructional evaluation, which occurs at the output point of instruction. In contrast, instructional evaluation involves a comparison of actual and intended learning outcomes in order to provide evidence of the effectiveness of learning but not of the curriculum.

In another work, Johnson (1978) attempts to demonstrate that curriculum development is a process and the curriculum its product. He further states that instructional planning, while also a process, results in the instructional plan or program. He argues that little progress can be made in curriculum unless such a distinction be maintained.

Curriculum as a Treatment

As previously indicated, external evaluation conceptualizes the curriculum as a treatment in order to evaluate its effects. This conception brings the assumption that a curriculum, once introduced into the evaluation situation,

remains relatively stable over long periods of time (Guba, 1978; Patton, 1980). Furthermore, this conception assumes that the curriculum is a causal agent whose effects on students' achievement can be certainly predicted and objectively observed (i.e., as measured by posttesting). Underlying this assumption is the paradigm's notion of "causality" which affirms that all phenomena "can be explained as the result (effect) of a cause that precedes the effect temporally or simultaneously with it" (Guba, 1983, p. 7). In A System of Logic Mill (1873) provides the logical procedure for establishing causality in controlled situations:

> If an instance in which the phenomenon under investigation occurs, and an instance in which it does not occur have every circumstance in common save one, that one occurring only in the former, the circumstance in which alone the two instances differ is the effect, or the cause, or an indispensable part of the cause of the phenomenon (p. 222).

This notion of causality has its origins in Aristotle's deterministic conception of the world. Aristotle conceives the world as an orderly whole where everything moves naturally toward a predetermined end. Events, according to him, are not haphazard or accidental but causally determined. He identifies four fundamental causes that determine an event: formal, material, efficient, and final (Riestra, 1970). Of these four causes, the final cause is the most important for Aristotle, as explained by Durant (1961):

> Of the varied causes which determine an event, the final cause, which determines the purpose, is

> the most decisive and important. The mistakes and futilities of nature are due to the inertia of matter resisting the forming force of purpose--hence the abortions and monsters that mar the panorama of life...everything is guided in a certain direction...the egg of the hen is internally designed or destined to become not a duck but a chick; the acorn becomes not a willow but an oak (p. 70).

Durant continues to explain that God is Aristotle's final cause:

> He is the final cause of nature, the drive and purpose of things, the form of the world; the principle of its life; the sum of its vital processes and powers; the inherent goal of its growth; the energizing entelechy of the whole. He is pure energy...He is not so much a person as a magnetic power (p. 71).

He further observes that this determinism is also central in Aristotle's philosophical psychology. According to Aristotle, "We cannot directly will to be different from what we are" (Durant, 1961, p. 72).

The adequacy of the paradigm's notion of causality for investigating physical phenomena has been challenged by Heissenberg's conception of the "uncertainty principle." This principle states that:

> ...it is impossible to assert in terms of the ordinary conventions of geometrical position and of motion that a particle (as an electron) is at the same time at a specific point and moving with a specified velocity (Webster's New Collegiate Dictionary, 1979, p. 1263).

This implies that we can never predict with certainty cause and effect relationships. Tranel (1981), in his article "A Lesson from the Physicists," argues that if the widely accepted notion of causality seems to be inapplicable in the "world of material substances, it is all the more inap-

propriate and misleading in the world of the individual, where measurement and predictability are inherently precluded" (p. 425). Similarly, Dobson, Dobson, and Koetting (1985), after illustrating the inapplicability of this notion of causality in the more measurement-oriented social sciences, express:

> If the model is not applicable in those social sciences that lend themselves to exactness, why have educators become so infatuated with the power of the cause-effect concept in the study of children? The ultimate thrust of the cause-effect model is prediction and control; emancipation and understanding are secondary concerns (p. 7).

They suggest the need for alternative concepts in the study of children in order "to enhance their educational living experiences" (p. 9). They personally support the view that new concepts "can be derived from aesthetic and ethical value systems" (p. 9).

Supporters of the prevalent conception of causality claim that cause and effect relationships can be "truly" established through the application of the "experimental method."

The Experimental Method. Basically, the experimental method consists of designing situations for observation in which the treatment is manipulated and extraneous variables are controlled.

Manipulation of the treatment means that the evaluator decides in advance the forms that the treatment will take and who will get which form (Gay, 1980). The different forms that the treatment may take are essentially presence

versus absence (treatment versus nontreatment), or presence of the treatment under consideration versus presence of a different treatment (e.g., treatment A versus treatment B). The application of the experimental method, typically, involves two groups, the treatment group, and the control group. The treatment group is exposed to the treatment under consideration, while the control group receives either the nontreatment form or the different treatment. Both groups are equated on all variables except the treatment variable.

The term "extraneous variables" is used to refer to those conditions not manipulated by the evaluator which might produce effects confounded with the effects of the treatment. Control refers to efforts to remove, or account for, the effects of extraneous variables. Uncontrolled extraneous variables are viewed as potential threats to the internal and external validity of the experiment. Internal validity refers to the condition that observed effects are a direct result of the manipulation of the treatment, not to some other factor; while external validity refers to the condition that results are generalizable or applicable to other situations.

Campbell and Stanley (1963) identify twelve classes of extraneous variables. These extraneous variables represent the effects of:

1. History - refers to the occurrence of any event that is not part of the evaluation situation.

2. Maturation - refers to developmental changes within the subjects.

3. Testing - refers to improved scores on a posttest resulting from subjects having taken a pretest.

4. Instrumentation - refers to changes in the calibration of evaluation instruments.

5. Statistical Regression - refers to the tendency of high scores to decrease toward the mean, or the tendency of low scores to increase toward the mean.

6. Differential Selection - refers to initial differences of the groups.

7. Mortality - refers to the drop out of subjects from the evaluation study.

8. Selection-Maturation Interaction - refers to the interaction of the initial differences and developmental changes.

9. Reactive Effects of Testing - refers to the subjects' sensitivity to the treatment as a result of the pretest.

10. Reactive Arrangements - refers to factors associated with the implementation of the treatment and the feelings and attitudes of participants.

11. Selection Treatment Interaction - refers to the possibility that the treatment and control groups are initially different.

12. Multiple-Treatment Interference - refers to the effects of prior treatments applied to the same subjects (pp. 5-6).

For controlling these extraneous variables, they strongly recommend the technique of randomization. This technique basically consists of selecting groups' participants or assigning the treatment by pure chance. It is based upon the assumption that through random selection and assignment differences between the treatment and control groups result only from the operation of probability or chance. These

differences are known as "sampling error" or "error variance," and their magnitude can be estimated through statistical techniques. If the difference between the performance of the treatment and control groups is too great to attribute to error variance, then it is presumed that the difference is probably due to the intervention of the treatment. Randomization is usually accomplished by flipping a coin, tossing a die, or using a table of a random numbers.

Curriculum Evaluation Models

The first formal application of the dominant paradigm in the evaluation of the curriculum may be credited to the physician/psychologist Joseph M. Rice. Rice (1897) evaluated scientifically the spelling curricula of 21 school districts in order to "prove" that the first step toward placing the elementary curriculum on a scientific basis must necessarily lie on predicting its learning outcomes. Rice measured the spelling achievement of 33,000 students using standardized tests.

Rice's work was furthered by the behavioral psychologist Edward L. Thorndike (DuBois, 1970). Thorndike recommended that once those outcomes were predicted they should be formulated in terms of students' behaviors. According to Thorndike, only behavioral outcomes were reliable indicators of learning (Seguel, 1966). He also recommended the use of behavioral outcomes as the standard for assessing the effectiveness of the curriculum. His recommendations

have been incorporated in the behavioral-objectives curriculum evaluation model that now dominates the field.

The Behavioral-Objectives Model

The educational psychologist Ralph W. Tyler, combining Thorndike's ideas, his rationale for developing achievement tests, and the product conception of the curriculum, formulated this model within the context of the Eight-Year Study (1933-1941). His model defines behavioral objectives as those specified behaviors to be exhibited by the student after a given period of instruction. A well-formulated behavioral objective, according to Tyler (1942), also includes the kind of situation in which those behaviors are to be elicited. Further elaborations of the concept of behavioral objectives (e.g., Mager, 1962; Vargas, 1972) include the notion of standard of performance. As Zais (1976) notes, the principle of operationalism comprises the theoretical base for the idea of behaviorizing educational objectives. He states:

> ...behavioral objectives are fundamentally neither new nor a creation of the educational establishment. The theoretical basis and much of the substance of behavioral objectives have been borrowed from physics and other disciplines of the scientific community from a long-established concept called "operationalism" (p. 311).

Zais goes on to define operationalism as the "process of consciously specifying those unambiguous observable indices that we agree imply the existence of the construct we wish to identify" (p. 312).

The basic purpose of evaluation, according to Tyler (1949), is to determine whether behavioral objectives have been achieved. He explains it best when he states:

> The process of evaluation is essentially the process of determining to what extent the educational objectives are actually being realized by the program of curriculum and instruction. However, since educational objectives are essentially changes in human beings, that is, the objectives aimed at to produce certain desirable changes in the behavior patterns of the student, then evaluation is the process for determining the degree to which these changes in behavior are actually taking place (pp. 105-106).

Tyler's evaluation procedure may be outlined as follows:

1. Derive curriculum objectives from three sources of data: (a) studies of the needs and interests of the learners, (b) studies of the society, and (c) reports of subject-matter experts.

2. Select those objectives that are consistent with the school philosophy and the psychology of learning.

3. Define those objectives in terms of measurable observable behaviors.

4. Identify situations in which the students can express the behaviors stipulated in the objectives.

5. Select or develop objective, reliable, and valid evaluation instruments.

6. Collect pre and posttest performance data relative to the objectives.

7. Determine the degree of congruence between these two sets of data. If a congruency between these two sets of data is found, then the curriculum is effective in meeting its purposes; that is, in changing the behavior patterns of students.

In his previous work, Tyler (1942) mentions at least three uses of the evaluation data: (1) for inferring hypothesis that can explain the weaknesses and strengths of

the curriculum, (2) for validating the underlying assumptions of the curriculum and (3) for deducing general principles that may provide an improved basis for guiding future curricular efforts.

Many curriculum evaluation scholars have furthered the work of Tyler by developing variations of his behavioral-objectives model. A few of them are Hammond (1973), Metfessel and Michael (1967), and Provus (1972). There are, however, others who have dissented from the Tyler model, and have proposed alternative conceptualizations, still within the dominant paradigm. Prominent among these alternatives are the Countenance, the CIPP, and the Goal-Free evaluation models. These three models are discussed below.

The Countenance Model

This model has been proposed by Stake (1967) as an alternative of what he considers to be the three major limitations of the Tylerian model: (1) the oversimplification of curriculum objectives, (2) the lack of concern for the teaching-learning process, and (3) the omission of judgmental data. Stake's Countenance model calls for attending to three phases of the curriculum: antecedents, transactions, and outcomes. Antecedents are the conditions prior to the teaching-learning process that may influence this process and, by implication, the outcomes. Transactions are defined as the "successions of engagements" that constitute

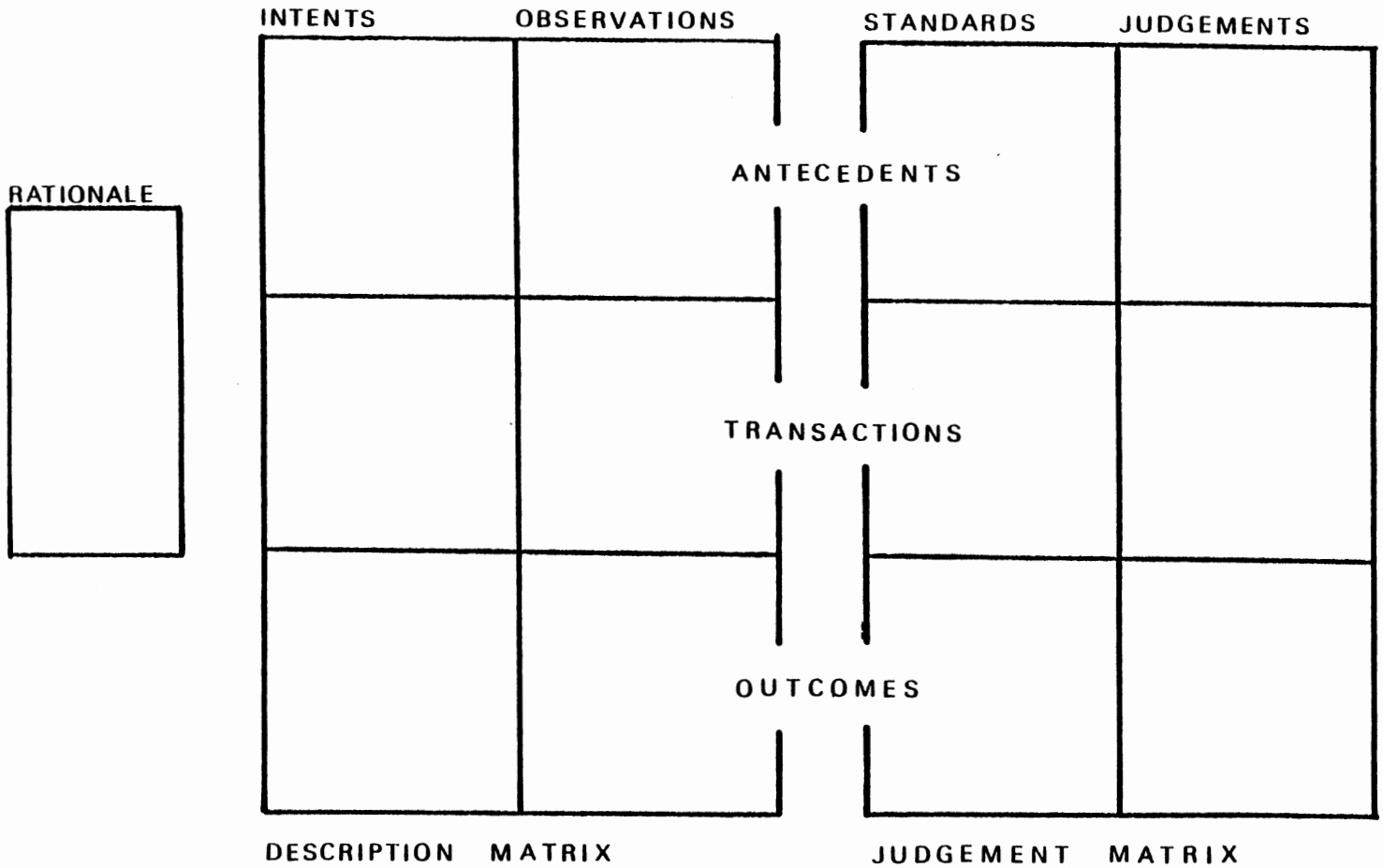
the teaching-learning process. Outcomes are considered to be the effects of the curriculum.

According to Stake, the general act of evaluation involves completing two data matrices: the description matrix, and the judgment matrix (Figure 3). Each matrix is divided into two columns, intents and observations for the description matrix, and standards and judgments for the judgment matrix. Both matrices are divided into three rows: antecedents, transactions, and outcomes. Basically, the process of evaluation involves the following steps:

1. Formulate intents for the antecedents, transactions, and outcomes.
2. Justify the intents in terms of some explicit rationale.
3. Set the standards for antecedents, transactions, and outcomes.
4. Collect data for the observations column of the description matrix.
5. Check the congruence between intents and observations data.
6. Check the discrepancies between these two sets of data against the standards.
7. Interpret the discrepancies.
8. Judge the merit of the curriculum.

In his countenance model, Stake suggests that evaluation should be systematic and scientific. Recently, Stake (1983) has moved away from the dominant natural science paradigm.

Several analyses of the Countenance model acknowledge its tremendous contributions to the concept of evaluation. As Guba and Lincoln (1981) note, this model has been the



Source: R. Stake. The countenance of educational evaluation, Teachers College Record, 1967.

Figure 3. Statements and Data to be Collected by the Evaluator in the Countenance Model

first to conceive the complete act of evaluation as involving both descriptions and judgments. However, the model has been severely criticized as inoperative (Lewy, 1973, 1977; Westbury, 1969).

The CIPP Model

The CIPP (Context-Input-Process-Product) model defines evaluation as "a process of delineating, obtaining, and providing useful information for judging decision alternatives" (Stufflebeam, 1973, p. 121). It focuses on four types of curricular decisions: (1) planning decisions, (2) structuring decisions, (3) implementing decisions, and (4) recycling decisions (Figure 4). Planning decisions are made to determine the intended-ends (goals). Structuring decisions are for the purposes of designing the intended-means (curriculum design). Implementing decisions are made on actual-means (the implemented curriculum) and recycling decisions are for the purpose of judging and reacting to actual-ends (curriculum attainments). Corresponding to the four types of decisions are four kinds of evaluation: (1) context evaluation, (2) input evaluation, (3) process evaluation, and (4) product evaluation.

Context evaluation serves the planning decisions by providing information relevant to the environment. Specifically, it provides information about unmet needs, unused opportunities, and desires and actual conditions. This kind of evaluation begins with a conceptual analysis in order to identify and define the limits of the domain to be

	INTENDED	ACTUAL
ENDS	PLANNING DECISIONS CONTEXT EVALUATION	RECYCLING DECISIONS PRODUCT EVALUATION
MEANS	STRUCTURING DECISIONS INPUT EVALUATION	IMPLEMENTING DECISIONS PROCESS EVALUATIONS

Source: D. L. Stufflebeam. An introduction to the PDK book. In B. R. Worthen and J. R. Sanders, Educational Evaluation: Theory and Practice, 1973.

Figure 4. Types of Decisions and Evaluations in the CIPP Evaluation Model

served by the evaluation. Its basic purpose is to provide a rationale for the planning decisions, i.e., the goals of the curriculum.

Input evaluation serves the structuring decisions. It assesses the curriculum design in terms of time, cost, constraints, relevance. Its intention is to provide an overall estimate of the potential of the design to meet the goals.

Process evaluation provides periodic feedback to curriculum implementers. This kind of evaluation has three specific purposes. First, to detect defects in the design. Second, to provide information for programmed decisions. And third, to maintain a record of the actual implementation. In general, process evaluation assists developers in making decisions for the modification and the improvement of the curriculum.

Product evaluation determines the overall effectiveness of the implemented curriculum. It is determined by relating the outcome information to the goals, and to the context, input, and process information. The purpose of product evaluation is to provide information for making decisions about the continuation, termination, or modification of the curriculum.

Stufflebeam (1973) provides a logical framework for designing a CIPP evaluation:

1. focusing the evaluation
2. collection of information
3. organization of information

4. analysis of information
5. reporting of information
6. administration of evaluation

He claims that the CIPP evaluation model has the potential for providing information that is valid, reliable, credible, timely, and pervasive. Like the Countenance model, the CIPP evaluation model has been criticized as difficult to operationalize (Lewy, 1973). In addition, it is argued that the CIPP model is highly expensive (Worthen and Sanders, 1973).

The Goal-Free Model

The Goal-Free evaluation (GFE) model has been conceptualized by Michael Scriven. Scriven (1974) maintains that the evaluator's knowledge of the goals intended to be achieved by a curriculum, which he calls the "rhetoric of intent" of the Tylerian approach, is a contaminated step in the evaluation process. He has arrived at this conclusion after experiencing many cases on which a curriculum had been judged as ineffective in achieving the intended goals while it had been rated high on the so-called "side effects" or unintended goals. In order to avoid contaminating the evaluation with the rhetoric of intent, Scriven proposes the use of the GFE approach. In GFE, the evaluator searches for all effects of the curriculum. There are no side effects to examine since data about all effects, and not only the intended effects, are of primary concern.

This type of evaluation is, according to Scriven, similar to Consumers' Union procedures:

> It seemed to me, in short, that a consideration and evaluation of goals was an unnecessary but also a possibly contaminating step. I began to work on an alternative approach--simply, the evaluation of actual effects against (typically) a profile of demonstrated needs in this region of education. (This is close to what Consumers' Union actually does.) I call this goal-free evaluation...(p. 35).

The GFE model supports the idea of using external evaluators (i.e., independent of the curriculum development process) for both the formative and summative roles. It argues that internal or staff evaluators, since they are too close to the developmental activity, tend to have "occupational tunnel vision" with respect to the effects of the curriculum; that is, a tendency to look mainly in the direction of intended goals (Scriven, 1974, p. 36).

This model also incorporates Scriven's (1967) early insistence that evaluators should not merely present information for administrators to use in formulating judgments about curricula, but the evaluators should arrive at and publicly report their own independent evaluative judgments. He asserts that the evaluators' main responsibility is to judge the relative merit of curricula since this is the principal goal of evaluation. He argues that the failure to recognize this goal has led to the "dilution" of the process of evaluation.

A recent analysis of Scriven's GFE model notes that although it is mainly theoretical it poses a challenge to the dominant position which holds that the act of evalua-

tion cannot proceed without the statement of intended goals (Guba and Lincoln, 1981).

Evaluation Methods

Curriculum evaluation relies upon the use of quantitative methods. The quantitative evaluation methodology has been developed to its greatest sophistication. Berk (1981) describes it as representing the "confluence of developments of measurement theory, research design, applied statistics, and computer technology" (p. 5). This section briefly discusses the major data collection instrument, required designs, and data analysis techniques.

Major Data-Collection Instrument

Achievement is the behavioral response most frequently measured in curriculum evaluations. Achievement data are collected via the administration of standardized achievement tests. A standardized achievement test is one that is: (1) constructed by measurement and subject-matter experts, (2) field-tested under uniform administration procedures, (3) validated to conform to the scientific criteria of reliability and validity, and (4) scored and interpreted using standard procedures (Gay, 1980). Standardized achievement tests are commercially developed. They are available for individual curriculum areas such as reading and math, and also in the form of comprehensive batteries which measure achievement in several different areas. The California Achievement Test Battery, for instance, is a

commonly used battery which measures achievement in reading, language, and arithmetic.

The vast majority of standardized achievement tests are norm-referenced tests (NRTs). Standardized NRTs have been previously administered to groups referred to as norm groups. Norm groups are selected with the intention of having a representative group with a wide spread of scores with respect to achievement (score variability). The scores of the norm groups are called norms, and they are represented in norm tables. The norms provide a standard of comparison and interpretation for other groups to whom standardized NRTs are administered.

NRTs interpret students scores in terms of their relative position with respect to the norms. Examples of indices or relative position are percentile ranks, standard scores, grade equivalents, and normal curve equivalents. As Berk (1981) notes, the evaluation of federally-funded curricula utilizes normal curve equivalents.

The vast majority of standardized achievement tests are norm-referenced tests (NRTs). Standardized NRTs have been previously administered to groups referred to as norm groups. Norm groups are selected with the intention of having a representative group with a wide spread of scores with respect to achievement (score variability). The scores of the norm groups are called norms, and they are represented in norm tables. The norms provide a standard of comparison and interpretation for other groups to whom standardized NRTs are administered.

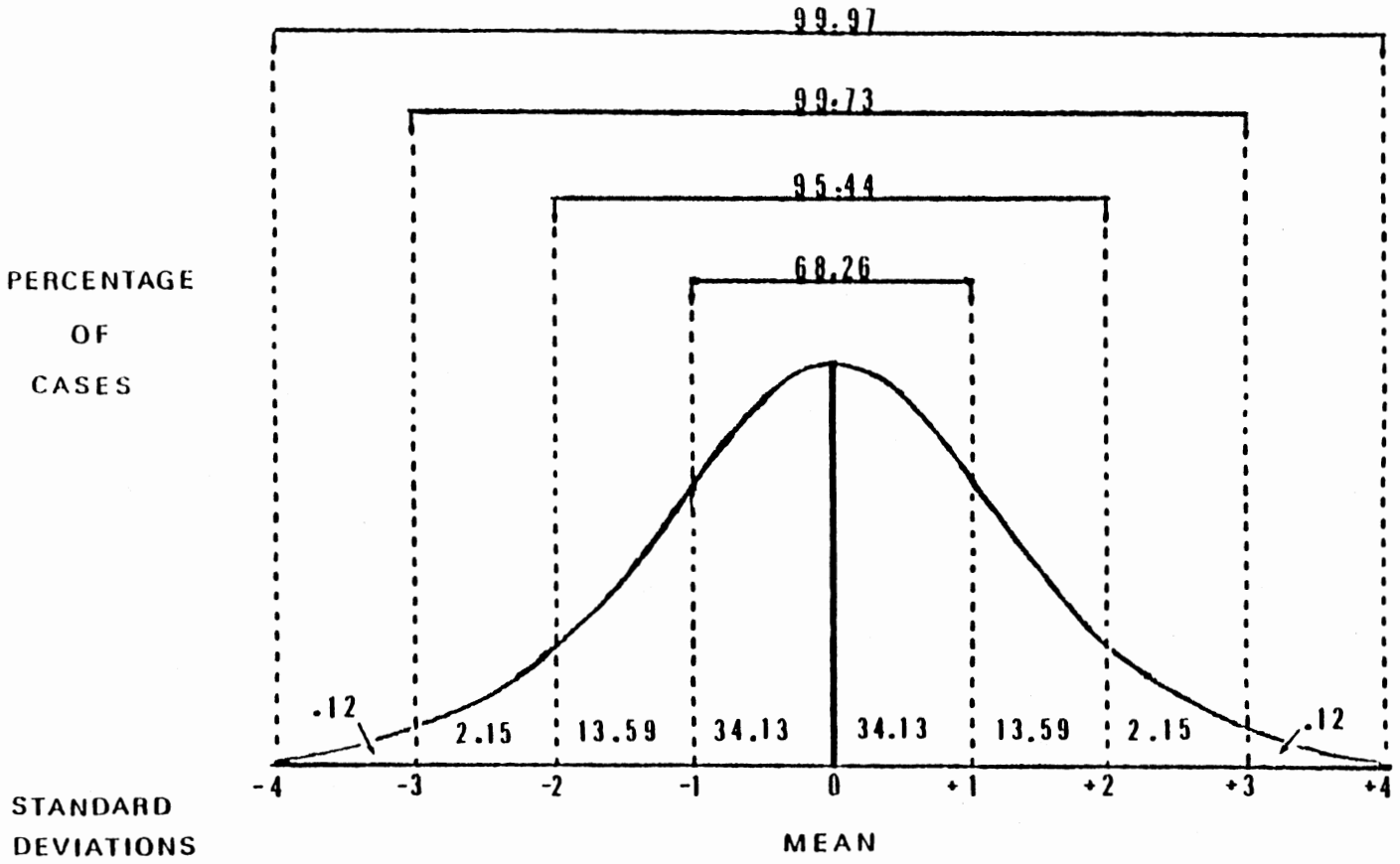
NRTs interpret students scores in terms of their relative position with respect to the norms. Examples of indices or relative position are percentile ranks, standard scores, grade equivalents, and normal curve equivalents. As Berk (1981) notes, the evaluation of federally-funded curricula utilizes normal curve equivalents.

Standardized NRTs are built on the assumption that achievement is distributed among students according to a Gaussian bell-shaped curve; i.e., the normal curve (Anastasi, 1976). Gay (1980) explains this assumption as follows:

> Norm-referenced standards are based on the assumption that measured traits involve normal curve properties.... The idea is that a measured trait, let us say math aptitude, exits in different amounts in different people. Some have a lot of it, some have a little of it, and most have some amount called an "average" amount (p. 140).

Based on this assumption, it is expected that the average group will obtain approximately 68% of the NRTs scores, and the above and below groups will contain 30% of the scores, that is 15% for each group.

According to Best (1970), the concept of the normal curve conforms to the "law of probability" formulated by the French mathematician/astronomer Abraham DeMoivre. This law explains the probable occurrence of certain events, and describes the fluctuations or chance errors of observation and measurement. The concept of standard deviation (an index of variability) in descriptive statistics is based on the law of probability and it is described with reference to the normal curve. Figure 5 depicts the relationship



Source: J. W. Best. Research in Education, 1970.

Figure 5. Percentage of Cases in a Normal Distribution Falling Within the Range of a Given Number of Standard Deviations

between a given number of standard deviations and the distribution of cases under the normal curve.

Standardized NRTs are often contrasted with locally developed criterion referenced tests (CRTs), which are used in the evaluation of individual students. CRTs report and interpret achievement scores in terms of an absolute standard. Typically, CRTs use as their interpretative frame of reference a specific content domain rather than a specific population of students. Glasser (1963) explains the advantage of CRTs as follows:

> Criterion-referenced measures indicate the content of the behavioral repertory, and the correspondence between what an individual does and the underlying continuum of achievement. Measures which assess student achievement in terms of a criterion standard thus provide information as to the degree of competence attained by a particular student which is independent of reference to the performance of others (as quoted in Nitko, 1984, p. 10).

According to Baker (1969), the intention of CRTs is to minimize variability and to emphasize mastery. The focus in on what is taught. Anastasi (1976) suggests that CRTs have a norm-referenced base since the choice of content is based on normative information. She states:

> A normative framework is implicit in all testing, regardless of how scores are expressed.... The very choice of content or skills to be measured is influenced by the examiner's knowledge of what can be expected from human organisms at a particular developmental or instructional stage. Such a choice presupposes information about what other persons have done in similar situations. Moreover, by imposing uniform cut-off scores on an ability continuum, mastery...does not thereby eliminate individual differences (p. 100).

In general, advocates of standardized NRTs do not view CRTs as being "evolved sufficiently to be considered a viable alternative to norm-referenced measures" (cf. Gay, 1980, p. 158).

Required Designs

Curriculum evaluation studies are based on research designs. There are two major classes of research designs, the single-variable designs and the factorial designs. The single-variable designs involve one treatment, which is intentionally varied (manipulated). These designs are categorized as pre-experimental designs, quasi-experimental designs, and true-experimental designs; depending upon the control they provide for sources of invalidity. Factorial designs are basically elaborations of the true experimental designs. The designs involve two or more treatments (at least one is manipulated). They are utilized for investigating the interaction of the treatments.

The evaluation of federally-funded curricula requires the use of one of these three single-variable designs: (1) the norm-referenced design, (2) the special regression design, or (3) the control group design. All three designs use norm-referenced standardized achievement test data; provide an estimate of expected no-treatment posttest performance; and report the results in normal curve equivalents (NCEs)

The Norm-Referenced Design. This design involves one group which is pretested, exposed to the treatment, and

posttested. Since the equivalent control group is missed, the norm group provides the control. Its initial procedures entail: (1) establishing the average standard score for the group at the time of pretesting, (2) converting the average standard score to a percentile, and (3) transforming the percentile to a normal curve equivalent (NCE). This NCE is taken as the expected no-treatment posttest performance. It is assumed that the expected NCE would remain stable over time without the treatment intervention. In order to determine the effectiveness of the treatment, the actual NCE of the treatment group is compared to the expected NCE. The treatment is judged to be effective if a significant difference between the two NCEs is found. By contrast, if no difference is found between the actual NCE and the expected NCE, the treatment is said to be ineffective; that is, the students performed as would be expected without the treatment intervention.

The norm-referenced design is a pre-experimental design. As the literature on research designs suggests, this type of design does not guarantee that the treatment accounts for the posttest differences. Campbell and Stanley (1963), for instance, identify at least five non-controlled variables that may provide rival explanations for the differences: (1) history, (2) maturation, (3) testing, (4) instrumentation, and (5) statistical regression. They maintain that the only way to insure that a treatment makes the difference is through the application of true experi-

mental designs since such designs control for almost all sources of invalidity.

The Special Regression Design. This is a quasi-experimental design; it is in essence a simulation of a true experiment; namely, the posttest-only control group design. The special regression design deals with those situations in which all members of an identified target population must participate in the treatment. This type of design selects the treatment participants on the basis of a cutting score on a quantified composite of qualifications. Specifically, any student whose score falls below the cutoff score will be assigned to the treatment; those who score above the cutoff score are assigned to the control group. At the appropriate time both groups are posttested.

The effectiveness of the treatment is determined by a statistical procedure called regression. It examines the regression line for predicting posttest performance. This predicted posttest performance is compared to the actual posttest performance. If a significant difference is found between them, the treatment is judged to be effective. For example, if an average posttest score of 10 was predicted for the group and the actual posttest average is 20, it is concluded that the treatment caused the 10 point difference.

The Control Group Design. This design typifies the true experiment. The control group design must satisfy three conditions. First, a single, well-defined population of students who has not been exposed to the treatment must

be identified. Second, two samples must be selected at random from this population. And third, the treatment must be randomly assigned to the different samples. That is, one sample must receive the treatment, and the other sample must provide the control (no-treatment). Both samples are pretested and posttested. According to Gay (1980), the combination of randomization, the pretesting, and the presence of the control group provides a control for all sources of invalidity.

In order to determine the effectiveness of the treatment, the posttest scores of both samples are analyzed either using directly a t-test or using analysis of covariance. Analysis of covariance adjusts posttest scores for initial differences on any variable, including pretest scores.

The use of this type of design in evaluation studies has provoked a heated controversy in the field. this design has been questioned on legal, ethical, philosophical, and feasibility grounds (Gay, 1980). Questions have also been raised over its claimed validity. In addition, it is argued that the use of experimental designs in evaluation conflicts with the notion of curriculum improvement. In this particular, Guba (1969) states:

> Perhaps the most damaging assertion that may be made about the application of conventional experimental design to evaluation situations is that such application conflicts with the principle that evaluation should facilitate the continuous improvement of a program. Experimental designs prevents rather than promotes changes in the treatments because, as has been noted, treatments cannot be altered if the data about differences

> between treatments are to be equivocal. Thus, the treatment must accommodate the evaluation design rather than vice versa (p. 35).

Patton (1975), in agreement with Guba, also comments:

> Because of a commitment to a single evaluation paradigm evaluators are frequently prepared to actually do everything in their power to stop program adaptation and improvement so as not to interfere with their research design. The deleterious effects this may have on the program itself by discouraging new developments and redefinitions...is considered a small sacrifice to be made in pursuit of a higher level of scientific knowledge (p. 33).

Data Analysis Techniques

Achievement test data are analyzed by statistical procedures. Two types of statistics are involved, descriptive and inferential.

Descriptive Statistics. This is the first step in data analysis. Descriptive statistical analysis is concerned with the organization and description of the data. There are four major types of descriptive statistics: (1) measures of central tendency, (2) measures of variability, (3) measures of relative position, and (4) measures of relationship.

The measures of central tendency are used to determine the average score of the group of scores. There are three major types of measures of central tendency: the mode, the median, and the mean. The mean ($\bar{X}$) is the preferred measure since it is the most stable. The $\bar{X}$ is the arithmetic average of the scores; it is computed by summing up all the scores ($\mathcal{E}X$) and dividing that total by the number

of scores (N). The formula is represented as $\bar{X} = \mathcal{E}X/N$. The $\bar{X}$ provides a single score or index that describes the performance of the entire group.

The measures of variability indicate how spread or disperse is the group of scores. The three measures of variability most frequently encountered are the range, the quartile deviation, and the standard deviation. The standard deviation (SD) is the preferred measure of variability in evaluation. Like the $\bar{X}$, the measure of central tendency which is its counterpart, the SD is considered the most stable measure of variability. These two measures provide a picture of how the distribution of the scores looks like. In a normal distribution, for instance, the $\bar{X}$ plus 3 SDs and the X minus 3 SDs encompass just about the 99% of the scores (Figure 6). This notion is expressed as follows: $\bar{X} \pm 3$ SD = 99% of the scores.

The measures of relative position indicate where a score is in relation to all other scores in the normal distribution. As suggested earlier, there are four major measures of relative position: the percentile ranks, the standard scores, grade equivalents, and normal curve equivalents. It was also noted that the evaluation of federally funded curricula uses the relative measures of normal curve equivalents (NCEs). The NCEs are standards which have been transformed to fit the normal curve (Anastasi, 1976; Gay, 1980). It is usually required a gain of 10 NCEs to judge a curriculum as effective.

The measures of relationship are used to find out to what degree the two sets of scores are related. The degree of relationship is expressed as a correlation coefficient. A correlation coefficient is a decimal number between .00 and ± 1.00. If the coefficient is near +1.00, the two sets of scores are positively correlated. If the coefficient approximates to -1.00, the two sets of scores are inversely related. Coefficients near the .00 indicate weak or no relationships.

Most correlational techniques are based on the assumption that the relationship being investigated is a linear one. The correlational technique most commonly used is the Pearson r. Like the $\bar{X}$ and the SD, the Pearson r is considered the most stable measure of relationship.

Inferential Statistics. The inferential statistical analysis is basically concerned with inferences or generalizations about the population based on the behavior of the treatment and control groups. For example, if a posttest difference between the means of the two groups was previously found the question of interest in inferential statistics is whether a similar difference exists in the population from which the two groups were selected.

Underlying the application of inferential statistics is the concept of the "null hypothesis." The null hypothesis is essentially the chance explanation for the difference found between the two sample means. It hypothesizes that there is no true difference in the population from which the samples were drawn, and that any difference found

in the samples is the result of chance or sampling error (Best, 1970; Gay, 1980, 1981). A null hypothesis might state:

> There is no significant difference between the mean math achievement of junior high school students who participate in curriculum A and the mean math achievement of junior high school students who participate in curriculum B.

The rejection or acceptance of a null hypothesis is based upon some level of significance as a criterion. In evaluation, the 5 percent level of significance (.05) is often used as the standard for rejection. Rejecting the null hypothesis at .05 level indicates that a difference in means as large as that found between the sample means would not likely have resulted from chance or sampling error in more than 5 out of 100 replications of the evaluation study. This suggests, 95 percent of the time, that the difference was due to the treatment rather than chance.

For testing the null hypothesis, a parametric test of significance is conducted. There are four major parametric tests: (1) the t-test, (2) simple analysis of variance (ANOVA), (3) analysis of covariance (ANCOVA), and (4) the Scheffé test. The selection of the appropriate test is determined by factors such as the measurement scale, the method of selection of the groups, the number of groups, and the number of treatments (Gay, 1980). The parametric tests make two assumptions about the nature of the samples:

1. The means of the samples will be normally distributed.
2. The mean of the sample means will approximate the mean of the population (Best, 1970).

In regard to the nature of the values of the population, the parametric tests assume that:

1. The population values are normally distributed or, if not, the nature of their distribution is known.

2. The population values have equal variances or the ratio of their variances is known (Best, 1970).

The results of the application of statistical analysis are interpreted in terms of the evaluation hypothesis.

As Eisner (1985b) notes, the use of statistics in evaluation is so pervasive that of the 47 studies published in the American Educational Research Journal in 1975 only one was non-statistical. It is believed that statistical studies are the epitome of good science (Patton, 1978).

This section has discussed the quantitative approach to evaluation. This approach, as noted, relies upon the use of standardized achievement tests, research designs, and statistical techniques. As Patton (1975) observes, the very dominance of this approach appears to have cut off the great majority of its practitioners from serious considerations of any alternative methodology. For instance, Campbell and Stanley (1963), in their widely used methodological primer, refer to this approach as "the only available route to cumulative progress" (p. 3). According to Patton (1975), practitioners of this dominant approach seem to be unaware that it is a philosophy upon which they stand. This underlying philosophy is examined in the next chapter.

CHAPTER III

THE POSITIVIST TRADITION

Introduction

The preceding chapter has presented the framework that the dominant paradigm provides for the field of curriculum evaluation. Underlying this framework is the philosophy of positivism. The essential attributes of positivism, according to the French positivist Auguste Comte (1798-1857), are those that correspond to the word "positive." By "positive" he means "relative, organic. precise, certain, useful, and real" (as translated by Bridges, 1957, p. 63). Comte also distinguishes positivism as the "only" philosophical world view which is "wholly disconnected with Theology" (Bridges, 1957, p. 100).

Philosophical positivism rests on the postulate of "scientism." It holds that, excepting knowledge of logical and mathematical system, science provides the model of the only kind of knowledge we can attain. This means that all that we can know is what we can observe, and it follows that everything we claim to know must be capable of empirical verification. Positivism thus denies the validity of nonempirical modes of knowing (e.g., "a priori" knowledge),

and equally denies the validity of theological and metaphysical knowledge. This latter position is explicitly stated by the eighteenth-century British philosopher David Hume (1711-1776), when he urges the destruction of theology and metaphysics:

> When we run over to libraries persuaded of these principles, what havoc must we make?... Take in hand any volume of divinity or school metaphysics...and let us ask: Does it contain any abstract reasoning concerning quantity or number? No. Does it contain any experimental reasoning concerning matter of fact and existence? No. Commit it then to the flames: For it contains nothing but sophistry and illusion (as quoted by Lavine, 1984, pp. 181-182).

According to Schroyer (1973), scientism has been defended in at least three different ways through the development of positivism. He notes, for instance, that Comte defends scientism by arguing that human society progresses to the extent that it utilizes the results of the natural sciences. While the English positivist John Stuart Mill defends this postulate by calling for the application to the scientific method in all spheres of inquiry. The exponents of logical positivism defend scientism, Schroyer observes, by claiming that the progress of human society is closely tied to the advancement of the sciences.

Critical theorists regard positivism as an "unreflective" philosophy. For example, Jurgen Habermas asserts that positivism has replaced reflection in epistemology for a concern with scientistic methodology. He refers to this as the "dissolution" of epistemology. His critique Knowledge and Human Interests (1971) is an attempt

to "recover" the "experience of reflection" in epistemology. He states:

> In following the process of the dissolution of epistemology, which has left the philosophy of science in its place, one makes one's way over abandonment stages of reflection. Retreading this path from the perspective that looks back toward the point of departure may help to recover the forgotten experience of reflection. That we disavow reflection is positivism (Habermas, 1971, p. vii).

Building upon Habermas, Michael Apple contends that the legitimation of this philosophy that "disavows significant critical self-reflection" in educational institutions has led educators to perceive their "style of scientific rationality as being interest free thereby contributing to an already strongly manipulative ethos of schooling" (Apple, 1975b p, 121).

Positivism is not a recent philosophy nor the conception of a single thinker. Its antecedents can be found in early thought. In ancient times, it is the Sophists and Epicureans who have been expressly singled out as representatives of positivism; in the Middle Ages, it is the Nominalists. For the modern period, Frondizi (1963) offers Otto Neurath's list of precursors corresponding respectively to England, France, and Germany: "Bacon, Hobbes, Locke, Hume, Bentham, J.S. Mill, Spencer; Descartes, Bayle, D'Alembart, Saint-Simon, Comte, Poincare; Leibniz, Bolzano, Mach" (p. 49). Charlton (1959), in reviewing positivism, discerns four major phases in its development since the seventeenth century to the present:

1. The British empiricist thinkers from Locke onwards and their French counterparts, the 'philosophies' from Bayle to Condillac and Condorcet;

2. The thinkers of the nineteenth century like Saint-Simon, Comte, Mill, Feuerbach, and Spencer;

3. The generation between 1880 and 1920, including Mach, Poincare, Duhem, and Russell; and

4. The present-day exponents of logical positivism (p. 4).

Charlton also observes that positivism has become the dominant philosophical world view of such countries as Austria, Germany, France, England, and the United States for the past eighty years.

Using Charlton's sketch, it can be said that positivism has proclaimed to be the philosophy of the natural as well as the social sciences since its second developmental phase. As far as 1848, for instance, Auguste Compte declared:

> Positivism has gradually taken possession of the preliminary sciences of Physics and Biology, and in these the old system no longer prevails. All that remained was to complete the range of its influence to the study of social phenomena (as translated by Bridges, 1957, p. 12).

In the social sciences in particular, Schroyer (1973) observes that positivism is committed to "provide more effective technical recommendations for piecemeal social engineering" (p. 118).

The remainder of this third chapter will be devoted to discussing the basic positivist assumptions concerning:

the nature of reality, the nature of subject-object relationship, the nature of truth statements, and the nature of value judgments.

Basic Assumptions

The Nature of Reality

Positivism rejects all theological, or metaphysical conceptions of the world. It argues that the world of the theologian, or of the metaphysician is an "ideal creation" independent from sense-experience that tells nothing about what is actually real. It then maintains that the world we do experience is the material world, and it is in this world that reality is to be found. This reality, according to positivism, is existentially "out there." It is one, i.e., a single and uniform reality; and it exists according to invariable natural laws. Positivism further asserts that we don't know the "essence" (or the ultimate nature) of this reality but rather its phenomena and their relationships. This knowledge of phenomena, however, is not absolute, but relative and conditional. Underlying this assertion is the assumption that the world is a determined world where its phenomena are invariably related in terms of cause and effect. This means that the world of phenomena is a world in which "if B has constantly been observed to follow A in certain conditions, then, given exactly the same conditions, we are entitled to suppose that B will

again follow A" (Charlton, 1959, p. 7). This ontological posture is best explained by J. S. Mill:

> We have no knowledge of anything but Phenomena.... We know not the essence, nor the real mode of production, of any fact, but only its relations to other facts in the way of succession or of similitude. These relations are constant; that is, always the same in the same circumstances. The constant resemblance which link phenomena together, and the constant sequences which unite them as antecedent and consequent, are termed their laws. The laws of phenomena are all we know respecting them. Their essential nature, and their ultimate causes...are unknown and inscrutable to us (quoted in Charlton, 1959, p. 7).

Prior to Mill, David Hume had argued that those who claim to know the ultimate nature of reality are either knaves or fools: fools because they don't understand that we are limited to sense perception in what we can know, and knaves insofar as knowing this limitation, they persuade us to follow a false philosophy (Lavine, 1984).

The phenomenal world of the positivist includes both physical and human phenomena. As Auguste Comte puts it, "...all events whatever, the events of our own personal and social life included, are always subject to natural relations of sequence and similitude..." (as translated by Bridges, 1957, p. 29). The purpose of positivist inquiry, then, is to explain phenomena and to formulate their laws. This inquiry, however, is confined to the "how." The "why," according to positivists, is a question to be left to the "imagination" of the theologians or to the "subtleties" of the metaphysicians (Charlton, 1959).

Guba and Lincoln (1981) specifically observe that positivist inquirers tend to see phenomena as "fragmentable into series of independent subsystems" (p. 56). These subsystems are often called "variables" and their relationships are expressed in the function $y = f(x)$ where x is defined as the variable to be manipulated and y is the variable on which the effect of the manipulation of x is to be determined. They note that the relationships of variables are principally explained for the purpose of predicting and controlling phenomena. Similar observations have been made by Guba and Lincoln (1983) and Guba (1983). Building upon these observations, Koetting (1984) distinguishes the positivist ontology as follows:

> For the positivist researcher, reality is a 'given.' It exists 'out there,' and can be divided into dependent and independent variables. These can be studied independently of each other. 'Inquiry can converge onto that reality until, finally, it can be predicted and controlled.' In other words, the world is seen as given, single, tangible, fragmentable, convergent (p. 9).

The Nature of Subject-Object Relationship

Positivism erects a separation, a dichotomy, between the individual human being (the subject) engaging in the activity of knowing reality (the object). Koetting (1984), basing on Egon Guba's discussion, puts it this way:

> The researcher maintains a distance between self and the object under investigation, 'neither disturbing it, or being disturbed by it' (p. 9).

This separation rests on the assumption that reality exists independently of the human mind. This means that reality

"is" whether or not the human mind is consciously aware of it or takes any interest in it (cf. Smith, 1983). Morris' (1961) description of what he calls a "common sense view of the ontology-epistemology relationship," though in a different context, is pertinent here:

> Reality is not an Absolute Mind thinking thoughts or an Infinite Self on which we are microcosms. Reality just is, without regard to human plans or purposes. It is, one might say, epistemologically neutral; to render it anthropomorphically we might say that reality is 'out there waiting to be known'; ontologically it is simply 'there' and awaits the onset of epistemological activity toward it. If anything can be called a common sense view of the ontology-epistemology relationship, certainly this can (p. 141).

Schroyer (1973) notes that positivism not only separates the knowing subject from the object of knowledge, but denies the "reflective participation" of the subject in the process of knowing. The subject confines himself to "copy" the elements of the independently existing object; knowing is merely a passive "picturing" of reality. He refers to this conception of knowing as "objectivism." Such a view of knowing, Schroyer points out, is rejected by critical philosophy. Critical philosophy places the human subject in an active role in the construction of knowledge and reality. In fact, it asserts that it is the human mind, human consciousness, who creates the object that is known through the act of self-reflection.

The positivist theory of knowledge, as the literature reveals, is fundamentally based on the sensationalism of John Locke and the skepticism of David Hume. John Locke, for instance, reduces knowledge to two basic questions.

The first question is: How do you know? If you can not answer Locke's question by showing that your knowledge rests upon sensory experience, observation of data, or experimentation with data, then you have no knowledge. He claims that sensory experience is the source and test of knowledge. The second question that Locke raises is: What kind of instrument is the human mind? He argues that "there is nothing in the mind except what was first in the senses" (as quoted in Durant, 1961, P. 256). The mind is at birth a "tabula rasa" or a "clean sheet" on which sense-experience writes upon it in a thousand ways (Durant, 1961); this writing by experience is all the mind can know (Lavine, 1984). Thus, the origin of our knowledge is in sensory experience through the mind receiving sensations from external objects.

David Hume, according to Lavine (1984), divides sensory experience into impressions and ideas. Impressions are our immediate sensations. Ideas are copies or images of impressions. Every idea has a corresponding impression from which it arises. As David Hume states:

> When I shut my eyes and think of my chamber the ideas I form are exact representations of the impressions I felt; nor is there any circumstances of the one which is not in the other...ideas and impressions appear always to correspond to each other (as quoted by Lavine, 1984, p. 153).

Hume argues that any idea which comes from no impression is worthless as knowledge and meaningless as an idea. Lavine refers to this argument as Hume's empiricist rule. He explains it in this way: "Where there is no impression,

there is no adequate idea. Where there is no impression, the idea is meaningless" (p. 155). He also notes that Hume basing on this rule proclaims that philosophical knowledge is not merely false but is meaningless since it derives from no immediate impression. Lavine considers David Hume as the "most destructive force in the history of Western philosophy" (p. 130).

The Nature of Truth Statements

Positivism rejects the "coherence" definition of truth advanced by idealist philosophers since this definition assumes that truth is mind-dependent. Positivism maintains that truth depends upon something which lies outside the mind. In view of this, positivism adheres firmly to the theory that our knowledge of truths is essentially a matter of correspondence.

Truth as Correspondence. In his work The Problems of Philosophy, Russell (1959), for instance, theorizes that truth consists in "some form of correspondence between belief and fact" (p. 121). A belief, according to him, is a "relation of believing or judging which relates a mind to several things other than itself" (p. 126). These several things are extrinsic to the mind: they belong to the independently existing world. They constitute what he calls a "fact." He then defines a "fact" as a "complex unity composed only of the objects of the belief" (p. 128). The correspondence between belief and this "complex unity" ensures truth. The absence of this correspondence entails

falsehood. Russell restates his correspondence theory of truth as follows:

> ...If we take such a belief as 'Otello believes that Desdemona loves Cassio', we will call Desdemona and Cassio the object-terms, and loving the object-relation. If there is a complex unity, Desdemona's love for Cassio', consisting of the object-terms related by the object-relation in the same order as they have in the belief, then this complex unity is called the fact corresponding to the belief. Thus a belief is true when there is a corresponding fact, and is false when there is no corresponding fact (p. 129).

He further asserts that the mind does not create truth. The mind creates beliefs, but once the beliefs are created, the mind cannot make them true. What makes a belief true is a fact, and this fact, according to him, does not in any way involve the mind of the person who has the belief. As he expresses:

> ...the condition of the truth of a belief is something not involving...any mind at all, but only the objects of the belief. A mind, which believes, believes truly when there is a corresponding complex not involving the mind, but only its objects (p.129).

Based upon his correspondence definition of truth, Russell draws a distinction between knowledge, error, and probable opinion. If what we firmly believe is true, we can call it knowledge, he maintains. But, in contrast, if what we firmly believe is not true, this is an error. However, if what we firmly believe is neither knowledge or error, then it is probable opinion. Russell argues that much of what commonly passes as knowledge is more or less probable opinion. He also contends that although a body of individually probable opinions may become more probable

than any one of them would be individually, probable opinions will never constitute "indubitable knowledge" (p. 140).

A pupil of Russell, the Austrian positivist Ludwig Wittgenstein (1899-1951) maintains that a statement about the world is true if it "pictures" actual facts for us (cf. Hartnack, 1965; Lavine, 1984). Wittgenstein refers to actual facts as present "states of affairs." And in any picture, he says, there must be a correspondence between the picture and the state of affairs it represents. Statements which fail to provide this picture are without any meaning at all, they are nonsensical. For Wittgenstein, only the statements of the natural sciences provide a picture of states of affairs. He argues that any kind of philosophizing must be directed to clarify the meaning of scientific statements. Philosophy, according to him, is not a doctrine but an activity. And he adds:

> The correct method in philosophy would really be the following: to say nothing except what can be said, i.e., propositions of natural science--i.e., something that has nothing to do with philosophy...(quoted in Lavine, 1984, p. 404)

According to White (1970), all versions of the Correspondence Theory of Truth rest on the principle of logical equivalence, which holds that "p is true if and only if p" (p. 98). This principle, as he indicates, was originally expressed in Aristotle's famous dictum "to say of what is that it is or of what is not that it is not, is true" (p. 102). He notes that since this principle suggest that there must be some additional item other than "what is

said" which makes "what is said true," most correspondence theorists insist that this item is a "fact" (p. 102). And "because the relation between what is said and this item is called a correspondence," their definitions of truth usually take the form of "To say that something is true is to say that there is a correspondence between it and a fact" (p. 102). White observes that correspondence theorists have found it difficult to explicate what they mean by "fact," "correspondence," and even "what is said."

Similar observations have been made by Deutsch (1979). But, he also notes that correspondence theorists are making efforts to elaborate an adequate explication of these terms since they feel that "there is something so undeniably right about the correspondence conception of truth" that it can not be abandoned (p. 81). Deutsch disagrees with this position and argues instead that what is needed is an alternative definition of truth. He proposes to conceptualize truth as the "achievement of rightness." He states his definition as follows:

> X is true when and only when it achieves rightness through the articulation of its own intentionality. X is perceived by Y to be true when Y recognizes that there is no correct alternative to X within the matrix of its presentation (p. 91).

By "X" Deutsch means anything that has the capacity to realize rightness, and "Y" refers to the person qualified to perceive that X is true. His definition rests on the assumptions that the world we experience is constituted by

particularity, and that truth is an inherent quality of particular things.

Kinds of Truths. For positivists, only two kinds of true statements exist: analytic and synthetic. Analytic statements are those tautological propositions found in the fields of logic and mathematics, while synthetic statements represent the empirical propositions of science. All other types of statements (e.g., the statements of metaphysics) are neither true nor false, but cognitively meaningless (Lavine, 1984; White, 1955). Hence, the concept of "meaning" (as defined by positivists) becomes the criterion for determining the truth of a statement. As Morris (1961) notes, positivist epistemologists utilize the method of "linguistic analysis" for establishing the meaning of propositional statements. This method, according to Morris, analyzes the linguistic relationship that a predicate of a proposition bears to its subject. He describes positivist epistemology as basically "an exercise in logic or an exercise in the analysis of the language we use to utter our ideas" (p. 173).

According to the positivist method, an analytic proposition is true because it is a "tautology," that is, its predicate is already contained in the meaning of the subject (Morris, 1961, p. 175). For example, the proposition "All husbands are married" is an analytic truth (cf. Popkin and Stroll, 1956, p. 182). The predicate "married" can be found in the meaning of the subject "husband" since the definition of a "husband" is that husbands are "married

males." This proposition is true by definition. It is also a logical truth. One cannot, without contradiction, deny the truth of this proposition. In other words, it is logically impossible to conceive of any circumstances in which somebody could be a husband and yet not be married.

The positivist Russell (1959) maintains that analytic propositions are general statements which assert logical relations of "universals." By "universals" he means abstractions. According to him, universals are "neither in space nor in time, neither material not mental" (p. 98). He also argues that universals do not "exist" but "subsist or have being, where being is opposed to existence as being timeless" (p. 100). For example, the mathematical proposition "Two and two are four" states a relation between the universal "two" and the universal "four." Russell further contends that the truths of analytic propositions are "self-evident," that is, they are incapable of proof. According to him, we can deductively infer other self-evident truths from analytic truths. He identifies three principles of formal logic assumed in such a deduction. They are the following:

1. The principle of identity: 'Whatever is, is.'

2. The principle of contradiction: 'Nothing can both be and not be.'

3. The principle of excluded middle: 'Everything must either be or not be.'

These three logical principles, according to Russell, are also self-evident truths.

One of the maximum exponents of twentieth-century positivism, Rudolf Carnap regards analytic propositions as "trivial" (cf. Popkin and Stroll, 1956). Carnap argues that from the truth of an analytic proposition we cannot infer that the items mentioned exist. For example, from the analytic truth that "All giants are giants" we cannot infer that there are any giants in the world. Thus, analytic propositions tell us nothing about the world of experience. They happen to be true merely by virtue of their logical form. Wittgenstein, as noted by Hartnack (1965), insists, however, that analytic propositions in a way do tell something about the empirical world. They tell us that the language we use to formulate the propositions and the world they claim to describe have the same logical structure. With regard to the language specifically, Popkin and Stroll (1956) note that Wittgenstein assumes that the language of a proposition resembles the structure of symbolic language of Principia Mathematica developed by Russell and Whitehead.

Since analytic truths make no assertions about the world of experience, they cannot be refuted by experience, maintain the positivists. For example, experience cannot refute the analytic truth that asserts that "two plus two equals four." This is a formal, abstract truth independently of any experience we might have (cf. Lavine, 1984). Thus , analytic propositions cannot establish any truth about existence. Yet they are useful, claim the positivists. They provide some guidance in "our empirical

search for knowledge" (Ayer, 1969, p. 42); they also supply us with those "transformation rules" that we need for applying to the "only genuine" propositions about the world, the synthetic propositions of science (cf. Pap, 1969, p. 60).

A synthetic proposition, as defined by positivists, is one whose predicate is not contained in the meaning of the subject; rather, its predicate has been attached to the subject on the basis of experience (Morris, 1961). For example, the proposition "All bachelors are introverts" is a synthetic one; its predicate "introvert" is not contained in the meaning of the subject "bachelor." According to positivists, the meaning of a synthetic proposition is to be found not in the words that make up the proposition, but in its method of empirical verification. As F. Waissman puts it:

> If there is no possible way to determine whether a statement is true then that statement has no meaning whatsoever. For the meaning of a statement is the method of its verification (quoted in Popper, 1961, p. 40).

That is, if you cannot describe what sorts of observations would lead you to verify the truth of your assertion, then, your assertion is meaningless; or, as Rudolf Carnap says, "your assertion is no assertion at all; it does not speak about anything; it is nothing but a series of empty words; it is simply without sense" (in White, 1955, p. 211). Thus, a synthetic proposition is meaningful if and only if it is empirically verifiable. This condition is often

referred to as the positivist "verifiability criterion of meaning."

In reviewing the criterion, Erwin (1970) notes that positivists make the claim that in advocating its use they are following the suggestion of the physicist Albert Einstein. He quotes, for instance, the leading positivist Moritz Schlick as saying: "All I am trying to do is to stick consistently to Einstein's position and to admit no exceptions from it" (p. 19). Erwin also observes that positivists often acknowledge Wittgenstein as providing the original version of this criterion, when he said in his work The Tractatus Logico-Philosophicus (1922): "In the verification of a proposition lies its sense.... If we do away with all means of verification we destroy the meaning" (p. 21). According to Quine (1969), this positivist criterion rests upon the "dogma of reductionism" initiated by Locke and Hume. He describes this dogma as the "belief that each meaningful statement is equivalent to some logical construct upon terms which refer to immediate experience" (p. 116). He notes that several positivists, for instance Carnap, embarked on the project to reduce scientific generalization into singular statements of immediate experience; but scientific statements "fell short of reduction not merely through sketchiness, but in principle" (p. 134). Quine indicates that Carnap, in his later writings, seems to have abandoned reductionism in its radical form.

It is contended that the criterion of verifiability is a "strategem" devised by positivists with the intention to

destroy metaphysical theories, e.g., Hegel's metaphysics. This criterion rejects metaphysical propositions as meaningless. The positivist A.J. Ayer, for instance, states:

> If a putative proposition fails to satisfy this principle [i.e., the criterion of verifiability], and is not a tautology, then I hold that it is metaphysical, and that being metaphysical, it is neither true nor false but literally senseless (quoted in Morris, 1961, p. 176).

The philosopher of science Karl R. Popper dissents in this particular from his positivist "friends" (as he calls them) by expressing:

> If by the words 'nonsensical' or 'meaningless' we wish to express...'not belonging to empirical science', then the characterization of metaphysics as meaningless nonsense would be trivial; for metaphysics has usually been defined as nonempirical. But of course, the positivists believe they can say much more about metaphysics than that some of its statements are non-empirical.

And he continues,

> The words 'meaningless' or 'nonsensical' convey, and are meant to convey, a derogatory evaluation; and there is no doubt that what the positivists really want to achieve is not so much a successful demarcation as the final overthrow and the annihilation of metaphysics (Popper, 1961, pp. 35-36).

Popper further contends that positivists "in their anxiety to annihilate metaphysics, annihilate natural science along with it" (p. 36). According to him, scientific theories, e.g., laws and prediction, are not conclusively verifiables. About this, Russell (1959), however, insists that the validity of scientific laws and predictions can be verified with the inductive principle of probability logic. He states this principle as follow:

> (a) When a thing of certain sort A has been found to be associated with a thing of certain other sort B, and has never been found dissociated from a thing of the sort B, the greater the number of cases in which A and B have been associated, the greater the probability is that they will be associated in a fresh case in which one of them is known to be present;
>
> (b) Under the same circumstances, a sufficient number of cases of association will make the probability of a fresh association nearly a certainty, and will make it approach certainty without limit (p. 66).

According to Russell, this principle is assumed in the formulation of all scientific theories, and its truth, like all logical truths, is self-evident.

The Nature of Value Judgments

Positivists categorize value judgments a "pseudo-assertions." They maintain that although value judgments have as similar grammatical form to the assertive propositions of science (i.e., to synthetic propositions), they are not assertions at all. Value judgments, they argue, are merely disguised ways of formulating prescriptions or commands. For instance, to say the "Stealing money is wrong" is to assert nothing, but to prescribe that "You ought not to steal money" or to command, e.g., "Do not steal money" (cf. White, 1970, pp. 57-65). According to positivists, since value judgments have no assertional sense, they lie completely outside disputes of truth and falsehood. In this regard, Rudolf Carnap writes:

> Most philosophers have been deceived by this form into thinking that a value statement is really an assertive proposition and must be either true or

> false. Therefore they give reasons for their own value statements and try to disprove those of their opponents. But actually a value statement is nothing else than a command in a misleading grammatical form. It may have effects upon the actions of men, and these effects may either be in accordance with our wishes or not; but it is neither true nor false. It does not assert anything and can neither be proved nor disproved (in White, 1955, p. 217).

Carnap further claims that since value judgments are neither true nor false, they are theoretically senseless. As he puts it, "The propositions of normative ethics, whether they have the form of rules or the form of value statements, have no theoretical sense, are not scientific propositions..." (in White, 1955, p. 218).

Positivists, as the literature reveals, also advance an "emotivist" thesis of values. This thesis postulates that value statements, unlike assertive propositions, have no representative or symbolic linguistic function. Rather, value statements are, like laughing, crying, and singing, expressive. They express our personal feelings, attitudes, desires, and emotions. This thesis is explained in C.K. Ogden and I.A. Richards' interpretation of the ethical concept "good." They jointly write:

> This concept, it is said, is the subject matter of real ethics. This peculiar ethical use of the word 'good' is, we suggest, a purely emotive one. When so used, the word stands for nothing whatsoever, and has no symbolic function. Thus, when we use it in the sentence 'This is good' we merely refer to 'this,' and the addition of 'is good' makes no difference to our reference. When, on the other hand, we say "this is red,' the addition of 'is red' to 'this' does symbolize an extension of our reference, namely to some other red thing. But 'is good' has no comparable symbolic function; it serves only as

> an emotive sign expressing our attitude to 'this' (quoted by Frondizi, 1963, p. 52).

And according to positivists, expressions of emotions, too, are neither true nor false because they do not assert anything about the world; they are cognitively or theoretically senseless. Therefore, they conclude, we assign them to the realm of metaphysics. And Rudolf Carnap, for instance, adds that David Hume was right when he said that "only the propositions of mathematics and empirical science have sense, and that all other propositions are without sense" (in White, 1955, p. 224).

As suggested by Bernstein (1978) and Frondizi (1963), this "emotivist" thesis goes beyond its original subjectivist version of the nature of values as to negate the very possibility of the existence of values. Now, within this new axiological "doctrine," writes, for instance, Frondizi,

> ...we do not say anything when we use words which have been and are basic in our daily, philosophic vocabulary, we don't say anything about the object, act or person to whom we attribute the quality of being 'good,' but instead we express only our own emotional state.

And furthermore,

> We do not confer value upon an object by means of our pleasure, desire or interest as the subjectivists whom we have examined maintain. We are, instead, committing the error of thinking that we are talking about an object when we are really expressing a psychic condition (1963, p. 52).

Frondizi also notes that positivists claim that the main reason for adopting this emotivist position is the "complete impossibility of finding any arguments to prove

that this or that has intrinsic values" (p. 106). He argues that this position is in error. Positivists, maintains Frondizi, can not find such arguments because they exclude elements which do not coincide with their theories and admit only those which are favorable.

In this regard, positivists respond that in philosophical positivism there is no place for unscientific expressions (cf. Kraft, 1953, p. 15). Positivism is based on rigorous scientific reasoning. Positivists postulate three basic requirements of scientific reasoning: (1) unambiguous clarity, (2) logical rigor, and (3) congent argumentation. They maintain that judgmental expressions do not meet these three requirements. According to Kraft (1953), judgments are like metaphysical speculations: "subjective, matters of opinion, unverifiable. Lacking universal validity, they are matters of personal conviction, but do not represent knowledge" (p. 193).

This section has discussed the philosophical posture adopted by positivists toward the nature of ontological reality, knowledge, truth, and values. This positivist philosophy dominates the curriculum evaluation field. During the late 1960s a new paradigm emerged in the field challenging this dominant philosophy. This emergent paradigm has been identified as the aesthetic paradigm. This paradigm is the concern of the next chapter.

CHAPTER IV

THE AESTHETIC PARADIGM: AN ALTERNATIVE

Introduction

Chapters II and III have explained the dominant natural science paradigm. Chapter II presented the conceptual and methodological framework that this paradigm provides for the field of curriculum evaluation. Under this framework the curriculum is conceived either as a product or a treatment, and quantitative methods are employed in evaluating its effects on students' achievements. Chapter III discussed some of the positivist assumptions on which this framework rests.

As indicated in the introduction of this study, this dominant paradigm has been challenged by three emergent paradigms. These new paradigms were identified as the legal, the anthropological, and the aesthetic, and a brief overview of each was provided. It was also indicated that this study has as its second purpose the examination of the aesthetic paradigm specifically, hoping that this examination will provide the stimulus for researching the other alternatives. The purpose of this chapter is to provide such an examination.

The first part of this examination focuses on the assumptions that the aesthetic paradigm offers as an alternative to the dominant positivistic assumptions. The second part explores some aesthetic views of curriculum, and the final part discusses Elliot Eisner's approach to evaluation. Eisner, as already mentioned, has become the leading advocate of the aesthetic alternative in the field of curriculum evaluation.

Philosophical Assumptions

The aesthetic alternative is an interpretive paradigm which is basically oriented toward obtaining an understanding of the created world. It is based on the philosophy of art. Art, as noted by Deutsch (1979) originates in religious life. When art stood so close to religion, it was regarded as magical and ritual. Song was prayer, drama was divine performance, dance was cult. During the period of the Renaissance art emerged as a distinct, autonomous field. According to Deutsch, when art achieves its autonomy from religion, the concept of "quality" becomes its fundamental category, and the notion of "meaning" its presentational content.

The aesthetic paradigm takes as its starting point for understanding the world the act of human experience. Experience, as defined by Eisner (1985a), is the consequences of the human beings' interaction with the qualities of the

world. The world is conceived here as an organic whole constituted of particular qualities of considerable richness and complexity. Experiencing the world is not a passive recognition of qualities, but an active-reactive exploration of the qualities and their interrelationships. It is an act of creation and re-creation. This creative process involves a total participation of all human capacities; e.g., sensory perception, cognition, intuition, imagination, will, emotion, and feeling (cf. Reid, 1964). It is argued that there is also a somatic participation in the experiential act. As indicated by Berleant (1970), the empathic theorists Vernon Lee and Theodor Lipps have observed how muscular movements are an integral part of the experience in such a way that there is an emulative physical participation in the response. For instance, they have noted how the human body tends to experience a physical imbalance when encountering the quality of imbalance. They also maintain that in experiencing the qualities of mass and space one becomes part of it.

Experience is also regarded as highly selective. Selectivity is seen as an indispensable condition to get hold of the flux of experience. It is argued that the function of selection is performed by the values and beliefs that the individual brings to the world. These values and beliefs also guide the creative process. John Dewey, in Art and Experience,stated:

> Every individual brings with him a way of seeing and feeling that in its interaction with old material creates something new, something previously not existing in experience (quoted in Pepper, 1945, p. 60).

According to the aesthetic paradigm, the positivist notion of objectivity is a myth that has been fostered by the Lockean idea of the empty mind. Eisner (1985a), particularly, believes that the very question of objectivity itself reflects naiveté. Objectivity, according to him, is an "epistemological impossibility" because an "empty mind sees nothing" (p. 185).

Human experience is described as unique, personal, individual, and private. As an attempt to express and communicate the experience, humans create a variety of forms. As Eisner (1985a) observes, the dominant natural science paradigm employs conventional forms such as numbers and propositions to represent the experience. According to him, conventional forms are a "surrogate" of an experience; i.e., they stand in the place of a referent (the experience). In conventions, the relationship between the form and the referent is arbitrary. In order to know the meaning of the form one must be able to imagine the referent. Unlike the dominant paradigm, the aesthetic paradigm communicates through artistically expressive forms; e.g., literature, music, poetry, painting, sculpture, and film. Eisner claims that art forms contain in themselves the expressive content to which they are related. They embody the experience the individual seeks to express. He also argues that the use of expressive

forms is crucial when we need to describe emotionally loaded situations. To use a form that leaves out the emotional content is to render a limited view of the situation which can lead to bias and distortions. In this particular, he expresses:

> After all, what can be more biased than emotionally eviscerted fact describing conditions or situations that are emotionally significant to those in the situations being described. Distortion can result not only from what is put in, but also from what is left out (p. 196).

One of the central interests of the aesthetic paradigm is to contribute to enhancing and expanding the quality of human experience. To pursue this interest, the paradigm attempts to interpret the meaning of the experience embodied in the artistic form. It is believed that the artform is essentially a sign or symbol for an experience that the creator "means" by it (cf. Berleant, 1970). The aesthetician Isenberg (1955) insists that the adoption of the aesthetic point of view involves a concern for meaning only, and not for observable facts. "I should think," he remarks to an audience,

> ...that all of you in this room had read the lines from Hyperion many times before and that few of you had ever asked yourselves whether you agreed with them--and this not from any slackness of attention but from the very fullness and fineness preoccupation with the meaning... To be preoccupied with the aesthetic object implies no disregard with the 'content' of the poem--only a disregard of one function of that content, namely, its relationship to observable fact (p. 398).

Isenberg argues that it is possible to interpret meaning without being concerned with observable referents since

what art seeks is understanding and not verification of facts.

According to Eisner (1985a), one way of understanding the meaning of an experience is through "empathy," that is, to "imaginatively participate in the experience of another" (p. 192). He contrasts this mode of understanding with the inferential approach in experimentation, which makes inferences from observable behavior to what is nonobservable. In the inferential approach there is no need for empathy since "observables are used in a kind of statistical fashion" in order to estimate the "probability that this behavior means one particular thing or another" (p. 192). Eisner agrees with the anthropologist Clifford Geertz that man is "suspended in webs of significance he himself has spun," and that the analysis of those webs requires not an "experimental science" but rather an "interpretive one in the search for meaning" (p. 192).

For interpreting the meaning of artistic forms that employ the medium of language, the aesthetic paradigm relies heavily on the hermeneutical approach. Hermeneutics is a dialectical method of interpretation in which there is a constant back and forth movement between parts and wholes with no absolute starting and ending points (Smith, 1983). It is based on the assumption that no part can be fully understood apart from the whole, and, conversely, that the whole cannot be experienced without understanding the parts. This method of interpretation also assumes that the interpretation and understanding of human experience cannot

be pursued in the absence of the context of activity. As Smith (1983) points out, although hermeneutics employs a diversified methodology, the unity of the approach is maintained by concentrating in the wholistic character of the human situation.

In the visual arts, the search for meaning has become so developed that it has evolved into the separate discipline of "iconography" (cf. Berleant, 1970). As noted by Hesenmueller (1979), the iconographer Erwin Panofsky has proposed the extension of his "iconological" method of interpreting meaning in painting to the literary arts. Panofsky's method identifies three types of interpretations for three levels of meaning in painting. The first interpretation is called "pre-iconographic." It is concerned with the recognition and understanding of the "primary or natural" level of meaning. This first level of meaning contains the visual and expressive content. The second type of interpretation is referred to as "iconographic." Its task is to decode "secondary or conventional" meanings articulated in images or icons. Primary and secondary meanings are conscious and decodable. The third type of interpretation is called "iconological." Iconology interprets what is unconscious and accessible only to subjective understanding, that is, the "intrinsic" level of meaning. Panofsky's method of interpretation conceives art as fundamentally communicative. It is based on the Hegelian assumption of the deep interconnectedness of all historical phenomena.

The task of interpretation is never complete. It is believed that the artform, like reality, has multiple levels of meaning, and any interpretation will reveal only a partial understanding. The aesthetic paradigm allows for divergent interpretations. As Eisner (1985a) suggests, the criterion of truth is irrelevant here. According to him,

> Truth implies singularity and monopoly. Meaning implies relativism and diversity. Truth is more closely wedded to consistency and logic, meaning to diverse interpretation and coherence (p. 198).

The interpretation of the meaning of the experience is followed by a judgment of its value. According to Irwin Edman, the quality of human experience is valued for its intensity and depth.

> Whatever experience may portend or signify, veil or reveal, it is irretrievably there. It may be intensified and heightened or dulled and obscured. It may remain brutal and dim and chaotic; it may become meaningful and clear and alive. For a moment in one aspect, for a lifetime in many, experience may achieve lucidity and vividness, intensity and depth. To effect such an intensification and clarification of experience is the province of art (quoted in Pepper, 1945, p. 57).

The more vivid the experience and the more extensive and rich its quality, the greater its value.

Some Aesthetic Views of Curriculum

As mentioned earlier, the aesthetic paradigm was first proposed in the more general field of curriculum. It emerged out of a reconceptualist orientation toward that field, an orientation concerned with finding new ways of thinking and talking about curriculum. According to Mazza

(1982), reconceptualists advocating the aesthetic perspective in curriculum expose some of the limitations of using (1) behavioral definitions of objectives and learning as central categories in curriculum theory, and (2) technical and instrumental rationalities for valuing curriculum experience. They propose a replacement of these limited views for the major concern of how curriculum can enable students to develop personal meaning.

Huebner's (1966) work is often acknowledged as the first explicit statement about the importance of viewing the curriculum aesthetically. Huebner argues that dominant conceptions of curriculum are inadequate in that they tend to tie the education process solely to the world of man's technique while ignoring his spiritual and transcendental nature. This inadequacy stems, according to Huebner, from an overdependency on a conception of educational values as "goals to be reached or behaviors to be learned" (p. 101). He contends that to value the curriculum because it "produces something or ends somewhere" is to turn it into a "commodity for a future state" (p. 107). Such a view denies the intrinsic value of curriculum experience. He councels curricularists to see the curriculum with "psychical distance." By this he means to see the curriculum as an aesthetic form apart from its instrumental function of achievement of goals. This mode of seeing would also allow curricularists to concentrate on the intrinsic aspects of the curriculum cleared from technical concerns.

Building upon Huebner's work, Mann (1969) proposes to talk about curriculum as if it were a "literary object." According to Mann, a curriculum is similar to a literary work in that both have a story, and a network of selections representing a universe of possibilities that constitutes an assertion of meaning or a symbolic commentary upon life. This meaning can be disclosed, he maintains, by critiquing the design or patterns of relatedness of the selections. But he points out that since the curriculum has many designs to be explained and thus many meanings to be disclosed, no single critique is exhaustive. He also maintains that in critiquing curricula, critics must be selective. In order for critics to choose among the many designs of the curriculum, he suggests they use as a criterion their own personal knowledge of ethical reality; that is, what they know about right and wrong or good and bad.

Mann argues that to talk about curriculum as a literary object also would reveal a peculiar characteristic of curriculum: its "unconditionedness." This quality, according to Mann, shows how inappropriate technological talk is in curriculum. He writes:

> If you turn this proposition around and look at it from the other side, it discloses something important--the unconditionedness of the curriculum... And this explains why technological talk cannot comprehend a curriculum... For technological talk is precisely talk about conditions, conditioning, and the conditioned. It is talk locked in a means-end cause-effect structure which cannot be bent to describe curriculum as unconditioned immediacy (p. 29).

He proposes as an alternative to technological talk his method of "curriculum criticism." His conception is grounded on the techniques of the "new criticism" in literature.

Mann's conception of curriculum criticism has been expanded by others. For instance, Willis (1978a) identifies three additional critical focus traditionally used in literary criticism which curriculum critics may attend to: the creator of the work, the universe, and the audience. Willis and Allen (1978) suggest that curriculum experience can be analyzed with the four structural elements used in analyzing literary experience: the text, the place of the text in history, the response of the reader, and the reader's psychological and biographical make-up. Kelly (1978) outlines a "rhetoric for the curriculum" consisting of "metaphor," "point of view," "plot," and "theme" (pp. 116-117). He proposes these four concepts as language tools for curriculum criticisms.

Drawing from the tradition of art criticism rather than literary criticism, Vallance (1978) also conceives the curriculum as a work of art. She identifies eight similarities between curriculum and works of art that she believes justify approaching curriculum from an artistic perspective. According to Vallance, both curricula and works of art are (1) products of human construction, (2) means of communication, (3) transformations of forms, (4) selections of the total realm of experience, (5) participants in an ongoing historical development,

(6) products of a problem-solving process, (7) intentional activities, and (8) subjects of critical appraisal. She proposes the language of art criticism as the tool for illuminating the kind of experience that a curriculum provides to the students. She also describes six rhetorical devices that are consistently used by art critics to communicate an experience more vividly: (1) selective emphasis, (2) similes and metaphors, (3) incidental comparison, (4) implied technique, (5) implied movement, and (6) logically unnecessary adjectives. Vallance argues that curricula not only produce effects but they are also "lived in." They color students' experiences and create a personal environment that is unique for each student. According to her, the dominant language in curriculum "with its predilection for identifying casual relationships" cannot account for this "personal lived-in quality of curriculum" (p. 144).

Speaking from their positions of art educators, Eisner (1985a) and Greene (1978) are more concerned with how the curriculum can be constructed so as to create those conditions which allow students to develop their own personal meaning. They urge giving the arts a central place in the content of curriculum. Eisner (1985a), for example, argues that the emphasis on the three Rs has created an "unbalanced" curriculum that has weakened rather than strengthened the quality of children's education. He insists that it is not possible to have any semblance of curriculum balance unless the content areas needed for such

balance be included in the curriculum. Balance in the curriculum, he argues, is not simply a plea for the equal representation of the arts in curriculum content, but rather "an imperative for helping students learn how to expand their modes of consciousness" (p. 128). Eisner's notion of balance in curriculum is "rooted in an understanding of the nature and scope of cognition" (p. 128). According to him, the arts, like the sciences and mathematics, are cognitive activities; they are one of the symbol systems that humans use in order to know. Symbol systems are also means through which consciousness is articulated. He maintains that the absence of artistic symbols in curriculum deprives students of the kinds of meaning that they can learn to create. He also states:

> If education has as one of its major aims the development of each child's ability to create meaning from experience, and if the construction of meaning requires the use of skills applied within a symbol system, then the absence of such systems within the curriculum is an impoverishment of the quality of education children receive (p. 128).

He suggests that such an education produces an "impoverished" mind.

According to Greene (1978), the un-aesthetic character of education derives from a "sense of social structures and explanatory systems pressing down on human beings and rendering them passive: gazers, not see-ers; hearers, not listeners" (p. 169). She suggests that the artistic-aesthetic perspective is a challenge to such a view, and

she insists that this perspective "ought to be reaffirmed" in the school curriculum (p. 172).

Greene visualizes the curriculum as "a number of provinces of meaning, each one associated with the kinds of experiences available to young people of different ages, with different biographies, and different locations in the social world" (1978, p. 174). One of these provinces of meaning, according to Greene, is the art world. She argues that perceptual encounters with this world bring people in touch with themselves.

Freene also maintains that there are certain works of art that have been deliberately created to provoke the kind of reflectiveness that the school curriculum should develop in the students. She refers to this reflectiveness, after Alfred Schutz, as "wide-awakeness." By "wide-awakeness" she means an awareness of the quest of meaning which has so much to do with feeling alive in the world. It is an awakening of consciousness "originating in an attitude of full attention to life and its requirements" (1978, p. 169). This attention is active, not passive; passive attention is the opposite of full awareness. She also points out that "wide-awakeness" contributes to the creation of the self, and awakens people to their freedom. For this reason, she believes that the arts ought to be central to any curriculum constructed today. She insists that educators must devise ways of integrating the arts to all levels of teaching. This needs to be done, she believes, "with a clear perception of what it means to

enable people to pay, from their own distinctive vantage points, 'full attention to life'" (p. 163).

Dobson, Dobson, and Koetting (1985) propose, as a preliminary step of creating those conditions reflecting aesthetic values, the strategy of "awakening consciousness through dialogue" (p. 12). They assert that dialogue is a reflective process for "dealing with human qualities (internal manifestations of beliefs and values) which people bring to the arena of human interaction" (p. 13). It is a kind of conversation that seeks a form of transaction on which the freedom of those involved is maximized. According to these authors, the purpose of the dialogue is to clarify thoughts. They agree with Paulo Freire that an educational encounter placing dialogue as its center begins the dialogical process. And in this process, they assert, individuals "reflect on their being" and through the "building of new structures and meanings they become aware [that] they are building themselves in the process" (p. 12).

In a different context, but with similar concerns, Koetting (1984) maintains that to enter into a dialogical relationship demands a strong commitment to human emancipation. He highlights Freire's (1971) six essential conditions required of those who enter into dialogue:

1. a profound love of men
2. humility
3. an intense faith in man (this is an <u>a priori</u> faith in the person)

4. trust (established through dialogue)
5. hope (rooted in the person's incompleteness, and recognition of that incompleteness; constant search)
6. critical thinking (p. 12).

According to Koetting, these conditions are "neither naive nor unworkable. They become, for subjects engaged in emancipatory praxis, a basic orientation to life" (p. 12).

Eisner's Approach to Evaluation

Over the past decade, Elliot W. Eisner has been working with the collaboration of some of his students at Stanford University on a qualitative approach to evaluation attuned with the aesthetic paradigm. This approach is rooted in his interest in the role of the arts in human expression and in the contributions they make to human understanding. Rather than emphasizing the outcomes of the curriculum, Eisner's approach to evaluation focuses on the educational context in which the curriculum is actively and creatively experienced as valuable--the classroom. It installs the processes of classroom life at its center, and it demands that those processes be observed over extended periods of time. As Eisner (1985a) explains, his interest on process evaluation stems from the realization, shared with many other evaluators, that if we want to improve the ability of students to perform paying attention solely to outcomes is insufficient or even misguided. It is important, as he puts it, "to see how the game is played

during practice sessions and how rehearsals go before the curtain rises" (p. 179).

Eisner's approach to evaluation is based on the work of those who inquire into the fine arts, the art critics. It requires what he calls "educational connoisseurship" and "educational criticism." He argues that connoisseurship and criticism represent two of the "modes through which we come to understand and express what we come to know" (p. 102). The remainder of this section is concerned with these conceptions.

Educational Connoisseurship

Eisner (1985a) argues that the major function of evaluation from an educational point of view is the improvement of the quality of classroom life. But this improvement will result not from evaluation's attempts to discover scientific laws that can be applied universally to classrooms, but rather from enabling teachers and others concerned with educational practice to improve their ability to see and reflect about what they do. According to Eisner, educational practice is a complicated affair filled with contingencies that are extremely difficult to predict, let alone control. In order to make the appreciation of such complexity possible, he proposes the art of educational connoisseurship. By appreciation here Eisner means an awareness and an understanding of its characteristics and qualities.

Educational connoisseurship is the act of knowledgeable perception. It is the result of having developed "an array of differentiated schemata" that enable connoisseurs to discriminate subtleties, and to discern qualities and relationships that others are less likely to see (p. 153). Eisner defines "schemata" as a set of values, theories, models, concepts, and ideas. These "schemata" also allow connoisseurs to distinguish the significant from the trivial and to place what they see in an intelligible context. He agrees with U. Neisser that "schemata" perform essentially a selective function in perception. They regulate not only how but what is seen. Eisner also argues that the absence of schemata in our symbolic repertoire creates a kind of "self-fulfilling prophecy." He explains:

> One can look without seeing, listen without hearing, eat without tasting, and touch without feeling.... What one does not have a schemata for, one is less able to experience.... Where schemata exits, the probability of experience is increased. In short, we tend to experience what we know how to find (pp. 151-152).

He also shares with the art historian E. R. Gombrich the view that there is no value-free mode of seeing: "artists do not paint what they can see they see what they can paint" (p. 152). Thus, to be an educational connoisseur is to know how and what to see.

According to Eisner, educational connoisseurship, like any art, is capable of development and refinement, and when developed to a high degree provides a level of awareness that makes intellectual clarity possible. Essentially, it

provides the content needed for reflection. In another text, Eisner (1985b) suggests several ways to develop high levels of connoisseurship regarding educational practice.

> What is involved in the development of educational connoisseurship is, first, the opportunity to attend to happenings of educational life in a focused, sensitive, and conscious way. Second, it requires the opportunity to compare such happenings, to discuss what one sees so that perceptions can be refined, to identify events not previously perceived, and to integrate and appraise what has been seen (p. 221).

The role of theory is also important in the cultivation of educational connoisseurship. He states:

> For the development of educational connoisseurship, an understanding of different social sciences, different theories of education, and a grasp of the history of education is not simply an intellectual ornament to be acquired within a graduate program but an essential working tool (p. 222).

But above all, the development of educational connoisseurship requires an intense desire "to perceive subtleties" and "to become a student of human behavior" (p. 220).

Eisner further explains that educational connoisseurship, as important as it is, is nevertheless a private act. It does not require a public disclosure of perceptions. However, educational connoisseurship is a necessary condition for doing useful educational criticism. As he explains:

> Connoisseurship, generally defined, is the art of appreciation. It is essential to criticism because without the ability to perceive what is subtle and important, criticism is likely to be superficial or even empty (p. 219).

In essence, educational connoisseurship "provides the fundamental core of realization that gives criticism its material" (p. 220).

Educational Criticism

Eisner (1985b) refers to educational criticism as the public side of educational connoisseurship. It is the act of publicly disclosing the qualities that connoisseurship perceives so that others can enter into the situation. The educational critics' function here is to serve as a "midwife to perception" (p. 217). They confront the difficult task of "lifting the veils that keep the eyes from seeing" (p. 217). Eisner agrees with John Dewey that the end of criticism is basically the "reeducation of perception" (p. 217).

Eisner notes that the qualities that educational criticism attempts to disclose are characteristically nondiscursive. In using language, however, criticism does not aspire to translate those qualities from one modality to another, but rather to create what Max Kozloff terms a "rendering." A "rendering" is essentially "an account of an experience, and never, as is sometimes supposed, a substitute for an experience" (p. 223). Eisner also indicates that the language of the critic resembles much that of the artist. It informs not by pointing to the qualities but rather by intimation. It "presents to our consciousness what the feeling of those qualities is" (p. 226). The language of critics and artists is a

presentational rather than representational language. He distinguishes these types of languages as follows:

> Representational symbols, the type used in conventional discourse, are like signposts; they point one toward the qualities but are not themselves intended to possess expressive qualities. 'Listen, listen to the bird' is a literal discursive expression, but 'Hark! hark! the lark!' contains an energy absent in the former. The former is representational; it directs our attention so that a certain kind of experience can be had. The latter presents us with a form that itself generates the excitement of the experience. The former is a conventional utterance, the later is poetic (p. 228).

He suggests that presentational language is the central vehicle for revealing the qualitative aspects of classroom life.

An educational criticism of a classroom typically takes the form of a written document whose aim is to help others to better see, understand, and appraise the quality of classroom practice. This document has three interconnected aspects: description, interpretation, and evaluation.

The descriptive aspect of educational criticism is essentially an attempt to render in language the significant qualities of the classroom situation. Eisner identifies three major focal points of critical description. One critical focal point is the "pervasive qualities," that is, those qualities characterizing the situation. Another focal point of critical description is the "component qualities." These are the particular qualities within the whole. A third focal point is what he refers to as the "underlying qualities." Here the

educational critic may ask: "What are the rules by which educational life in this classroom operates?" (p. 231). The critical description of underlying qualities is fairly "thin," that is, it renders the qualities without getting very deeply into what they signify or mean. This task is pursued in critical interpretation. Eisner maintains that in order to perceive those qualities two conditions must be present: the qualities must be in the situation in the first place, and the educational critic must have the ability to note their presence.

Eisner regards the descriptive aspect of educational criticism as making the most artistic demands on the educational critic. The critic's task is to provide a rich and vivid rendering that allows others to participate vicariously in the qualities that they have not experienced directly. It is here, insists Eisner, that the "critic's verbal magic must be most acute" (p. 231). The critic must be able to capture the perceived qualities "through the possibilities within words" (p. 232). In this process, according to Eisner, the use of metaphor is centrally important. As he explains:

> In this process of transformation, metaphor is, of course, a centrally important device. Metaphor breaks the bonds of conventional usage to exploit the power of connotation and analogy. It capitalizes on surprise by putting meanings into new combinations and through such combinations awakens our senses. Metaphor is the arch enemy of the stock response (p. 226).

He also argues that "nothing is more precise" in critical description than the use of metaphor (p. 227).

Eisner remarks that critical description is not a chronicle of happenings, but a creative construction of the situation; it is human construction.

The second aspect, critical interpretation, represents an effort to empathetically understand the meaning of what has been perceived and vividly described. Critical interpretation is guided by the general question: "What does the situation mean to those involved?" (p. 233). The interpretation of meaning requires an understanding of contextual conditions. According to Eisner, the role of interpretation in educational criticism is closely related to the notion of "thick description" used by the anthropologists Gilbert Ryle and Clifford Geertz. Thick description aims at describing the significance of human events or behavior as they occur in a cultural network saturated with meaning.

The role of theory is central in critical interpretation. Theory provides the conceptual maps and interpretive tools for understanding meanings. He recommends educational critics to ground their interpretations specifically on theoretical ideas of the social sciences, and of the history and philosophy of education. Critical interpretation also requires, according to Eisner, the judicious and informed use of the knowledge gained through training, and experience in classrooms.

The evaluative aspect of educational criticism attempts to appraise the educational value of what has been perceived, described, and interpreted. The major function

of the educational critic here is to apply educational criteria so that judgments are made on educational grounds. In the performance of this task, knowledge of the history and philosophy of education is crucial. The former provides the context necessary for purposes and comparisons and the latter the theories from which grounded value judgments can be made.

According to Eisner, it is the evaluative aspect of educational criticism that most sharply distinguish the work of educational critics from that of the anthropologists. Anthropologists, for instance, aspire to appear invisible within the culture they study. In addition, they have no obligation to appraise the value of a culture since they have no professional mission to change or improve it. Their interest is with understanding. Educational critics, however, are neither neutral observers nor disinterested interpreters. Furthermore, they are not only concerned with understanding but also with improving the educational situation. In this regard they have the professional obligation to judge.

Eisner notes that one of the issues that is most frequently raised about educational criticism deals with objectivity. In response to this issue, Eisner redefines the notion of "objectivity" as a "function of intersubjective agreement among a community of believers" (p. 241). In this view the question of whether criticism is really true does not proceed. Questions, for example, about usefulness and relevance are more adequate. Within this view,

then, educational criticism receives consensual validation through the criteria of "structural corroboration" and "referential adequacy" (p. 241).

Structural corroboration seeks to establish the extent to which criticism forms a coherent whole. It seeks to determine if the pieces of evidence hold together and make sense. Referential adequacy is determined by checking the relationship between what the critic has written with the subject matter of criticism. If the criticism is useful, the reader should be able to experience the situation in a new, more adequate way.

Beyond the issue of objectivity, Eisner explains that the need for unanimity among critics is not characteristic of educational criticism. It recognizes that complex situations such as classrooms are as "multilayered" as works of art, and different critics will find different qualities to describe, interpret, and evaluate. Educational criticism, as Eisner points out, does not seek a single definitive criticism but rather criticism that is useful. He writes:

> The cultivation of such productive diversity is a virtue, not a vice. As in education itself, we do not seek to create an army marching in step to the same tune but individuals who follow their own drummer as long as the beat is interesting (1985b, p. 244).

For Eisner, the major virtue of his proposed approach is that it expands our understanding of how we come to know, and as a consequence it makes new avenues for evaluation possible. He believes that even if this

approach does not succeed in becoming the major mode of evaluation, that contribution may be enough to have made the effort worthwhile.

The acceptance of Eisner's approach in particular and the aesthetic paradigm in general has significant implications for the curriculum evaluation field. These implications are discussed in the coming chapter.

CHAPTER V

IMPLICATIONS, CONCLUSIONS, AND RECOMMENDATIONS

Introduction

The previous three chapters were devoted to fulfill the basic purposes of this study. Chapters II and III examined the natural science paradigm, a paradigm which was initially identified as dominating the field of curriculum evaluation. A proposed alternative to this dominant paradigm, namely, the aesthetic paradigm, was then presented in Chapter IV. The information thus far obtained provides for answering the guiding questions of this study concerning the philosophical, conceptual and methodological tools that these paradigms provide for practicing curriculum evaluation activities.

The present chapter completes this study of paradigms. It is divided in three additional sections. The first of these sections discusses some of the implications of the aesthetic alternative for the field of curriculum evaluation. The second is a concluding section. Finally, the closing section of this study offers some recommendations for future studies of paradigms in curriculum evaluation.

Implications

Proponents of the aesthetic paradigm insist that this is the paradigm of choice. They make a number of educational claims on its behalf. Eisner (1985a), for example, maintains that this paradigm offers the tools for illuminating the richness and complexity that constitutes life in classrooms. Jenkins and O'Toole (1978) claim that it is of special relevance for understanding the essential meanings of the hidden curriculum. Willis (1978a) insist that this alternative must be chosen in order to assure a full and comprehensive view of educational environments. But what is the significance of choosing this alternative for guiding the field of curriculum evaluation? What, in other words, is implied in this choice? In choosing the aesthetic paradigm, the following seven implications are evident.

First, the aesthetic paradigm implies that the curriculum can only be evaluated in relation to the context in which it functions, that is, the total classroom situation. If this notion is accepted in the field, certain significant consequences follow. One of these is the rejection of the experimental approach to curriculum evaluation. This approach, as previously noted, tends to isolate the curriculum being evaluated from contextual conditions. This is usually done through the application of the strategy of randomization and statistical techniques, or by imposing the artificiality of the laboratory setting. The laboratory, as Guba and Lincoln (1981) point out, is the "essence of the

context-free environment" (p. 74). Another consequence is the abandonment of the behavioral evaluation model. Behaviorists take a reductionist stance; that is, they reduce the curriculum to a set of specific objectives which are then evaluated independently of the whole. Adherents of the aesthetic alternative, on the contrary, assume a wholistic stance. In fact, Willis (1978b) makes the claim that the phenomenon called curriculum defined in its fullest dimension is itself the classroom situation taken as a whole. A third effect is that the emphasis on product evaluation must be replaced for a process orientation. To be concerned with process demands evaluators to spend substantial periods of time in classrooms. Process evaluation is based on extensive observation. A product evaluation is a quick procedure which usually takes 15 to 20 minutes (Eisner, 1985b). It is often referred to as the "hit-and-run" approach.

Second, the aesthetic paradigm implies that the evaluator actively participates in the evaluation situation. This implication involves a rejection of the positivist assumption of passivity. Positivism, as noted earlier, assigns human beings the role of passive observers who confine themselves to copy the elements of a given reality. The aesthetic alternative, on the contrary, places human beings in the role of active participants. In addition, it replaces the notion of given-ness for a constructivist conception of reality. According to constructivists, human beings create

their own reality through a construction process analogous to the creative process of the artist (Donmoyer, 1981).

Third, the aesthetic paradigm implies that the evaluator should have practical experience in classrooms. As Eisner (1985b) observes, since the 1960s much curriculum evaluation work has been done by outsiders of the field, particularly behavioral scientists. He suggests that the principal motive of this group is not so much to change classroom practices as it is to advance their research methodologies. He maintains that it is the function of evaluation to change classrooms. Since the task of the evaluator, according to Eisner, is to make judgments about what needs to be changed, he/she must have a great deal of knowledge and experience with classroom practices; that is, he/she must be a "connoisseur." This notion suggests that evaluation should be done by internal rather than external evaluators.

Fourth, the aesthetic paradigm implies that classroom observations are value-dependent. This means that the classroom situation can not be objectively described. Descriptive statements about classrooms are value-laden. A significant effect of acknowledging the presence of values in descriptions involves a rejection to the positivist claim that assertive statements are value-free. This fourth implication is important to the field because it will require evaluators not only to make explicit their values but also to justify their value choices.

Fifth, the aesthetic paradigm implies that curriculum evaluation is essentially a judgmental activity. This means that evaluators, in addition to describing the classroom situation, must render judgments of value. According to Eisner (1985b), since there are many conceptions of what constitutes educational virtue, evaluative judgments would differ one from the other; but the need to make judgments is inevitable because evaluation is, after all, a normative discipline. He also maintains that the fact that judgments differ is not necessarily a liability in evaluation; it could be a strength. For example, it could open up the kind of discussion that educational practice should but does not now receive. According to him, educators for too long have approached educational problems as if simple and universal solutions existed. Discussion of values would encourage educators to seek more diverse solutions to educational problems. This idea of rendering judgments is rejected by adherents of the dominant paradigm. As mentioned earlier, they hold that judgmental statements are based on value terms that are primarily the expression of emotions rather than assertions of anything. They are, therefore, cognitively meaningless.

Sixth, the aesthetic paradigm also implies that value judgments are relative. Relativism holds that there are no objective standards on which any value claim can be based; the very notion of standards is value-dependent. The basis of judgments is then relative, differing according to

events, circumstances, people, etc. This relativistic stance can be noted, for instance, when Eisner (1985b) argues that

> The differences in basic assumptions among Freudiens, Rogerians, Skinnereans, Heiderians, Eriksonians, Piagetians, and the like are not resolvable through science. The fundamental theoretical structures through which each defines psychological reality differ, and there is no critical test that will resolve the truth or falsity of their respective belief systems. Each has a community of believers who reaffirm the beliefs of those working within the system (p. 241).

He then maintains that to claim that a belief system is true is to embrace a hopeless correspondence theory of truth. In regard to judgments of value, the leading positivist Carnap, for instance, argues that the fact that they can not be proved or disproved through science demonstrates that they are not scientific statements; value judgments are simply metaphysical expressions without any sense at all (in White, 1955).

Finally, the aesthetic paradigm implies that descriptive as well as normative statements in curriculum evaluation contain both emotional and cognitive components. As previously indicated, the dominant position holds that only descriptions of what is are cognitive statements. These statements can be verified empirically. But the statements of normative ethics have no empirical referent since they are non-cognitive. They are simply emotive expressions. According to the aesthetic perspective, the separation of emotions and cognitions misconstrues the nature of human

experience. It then postulates that human experience constitutes a unity. As Eisner (1985a) expresses it:

> This orientation to knowledge embraces an epistemology that rejects the positivistic view which holds that only formal propositions can, in principle, provide knowledge. It rejects the view that affect and cognition are independent spheres of human experience (p. 198).

In this particular, the aesthetician Berleant (1970) contends that if the unity of experience is recognized, a rejection to the emotivist thesis of values advanced by positivists follows. As a consequence, the status of aesthetic inquiry will be restored. According to him, positivists by assigning a non-cognitive dimension to the evaluational content of normative judgments have contributed to the exclusion of aesthetic inquiry from the realm of knowledge. Like Eisner (1985a), Berleant maintains that aesthetic inquiry is a legitimate cognitive activity.

Conclusions

This study was introduced by asserting that the field of curriculum evaluation is at present in a crisis state. This assertion was further supported by identifying four major signs indicative of this crisis. One of these signs, the emergence of competing paradigms, provided guidance for articulation of the basic purposes of this study. First, this study examined selected philosophical, conceptual and methodological aspects of the paradigm presently dominating the curriculum evaluation field. This dominant paradigm, as

indicated, comes from the positivist tradition of the natural sciences. As a second purpose, a proposed alternative to this dominant paradigm, namely, the aesthetic paradigm, was then explained. In addition to this examination, this study explored the major implications of accepting the aesthetic alternative for curriculum evaluation.

In relation to the paradigmatic aspects here examined, four conclusions are presented:

First, the natural science and the aesthetic paradigms are based on differing conceptions of the world. In fact, so different are the worlds they define that one gets the immediate impression that there is no means of comparing these conceptions at all. It is as though their advocates were speaking of two separate worlds: the scientific and the aesthetic worlds. For example, scientists speak of a world governed by invariable natural laws where its phenomena are linearly related in terms of cause and effects. Artists, on the other hand, speak of a world constituted of particular qualities of inexhaustible complexity and multiple realms of meaning. No wonder, the aesthetician Ernest Cassirer figuratively expressed in An Essay of Man (1953) that "...art and science move in entirely different planes, they cannot ... thwart one another" (quoted in Eisner, 1985b, p. 225). But since art and science are in competition, they indeed have established a point of comparison. That is, artists and scientists are

expressing their views about a common world. And these views of the world differ one from another.

Second, these two paradigms are based on opposing philosophical assumptions. Ontologically, the dominant natural science paradigm assumes that reality is a given. It exists "out there" independently of the human mind; that is, it is objective. The aesthetic paradigm, however, postulates that reality is a human construction. It is constructed and re-constructed in the process of interpretation. These interpretations of reality are viewed as diverging since it is believed that they are influenced by personal values and beliefs.

At the level of epistemology, the dominant paradigm conceives knowing as a passive picturing of facts. According to the positivist Wittgenstein, any knowledge claim is true if there is a correspondence between the picture and the facts it intends to represent. This correspondence is verified through the application of the scientific method. The aesthetic paradigm rejects this assumption of passivity. It acknowledges the active involvement and response of the participant in the process of knowing. Knowing is described here as the exploration of subtle particulars and their interrelationships. This alternative also replaces the concept of "truth" for the notion of "meaning". According to Eisner (1985a), truth implies singularity and consistency; while meaning implies diversity and coherence.

In relation to axiology, the dominant paradigm assumes that only those statements couched in scientific and numerical statements are truth assertions about the world. These terms are considered to be value-free. The aesthetic paradigm holds that every statement is valuational. It posits that there is no value-free mode of seeing. Gombrich's expression is often cited as representing this axiological posture: "artists do not paint what they can see; they see what they can paint" (Eisner, 1985a, p. 152).

Third, these paradigms offer evaluators a distinct definition of the entity to be evaluated, that is, the curriculum. For instance, a practitioner of the dominant paradigm works within a behavioral definition of curriculum. The curriculum is defined here as consisting of a series of intended outcomes expressed in terms of measurable, observable behaviors. The function of the curriculum, according to Tyler (1949), is to change the behavior patterns of the students. The evaluation of the curriculum involves an assessment of effectiveness. The curriculum is considered effective if a congruence exists between its intended outcomes and the behavior exhibited by the students at the end point of instruction. This congruence is determined by statistically analyzing the results of standardized achievement tests. Under the aesthetic paradigm, the curriculum is conceived as analogous to the works of artists. Works of art, according to Greene (1978), are deliberately created to provoke awareness and critical reflection. It is also believed

that the curriculum, like a work of art, has multiple levels of meaning that can be experienced and disclosed. The process of evaluating the curriculum is similar to criticism in the arts. The curriculum evaluator is to attune him or herself to curricular activity much as an art critic would experience and appraise a painting or a film. According to Vallance (1978), the goal of the evaluator is to render a judgment of the quality of experience that the curriculum provides for the students.

And finally, these paradigms provide their adherents incompatible modes of practicing curriculum evaluation. For example, evaluators working within the dominant paradigm employ scientific approaches to curriculum evaluation; while evaluators adhering to the aesthetic alternative paradigm rely upon the use of artistic approaches. The present study reveals that there are six major dimensions on which these approaches to curriculum evaluation differ.

One difference resides on the focal point. The scientific mode, for instance, focuses on observable behavior. The behavior of sample groups are statistically analyzed in order to make inference about the population from which the samples were selected. The question of interest here is whether a difference as that found in the samples also exists in the population. The artistic mode, in contrast, is concerned with what is nonobservable. It focuses on the meaning dimension. The major interest here is in understanding the meaning of human experience.

A second difference between these two modes is in emphasis. The scientific approach to evaluation is outcome-oriented. It emphasizes the measurement of curriculum outcomes. The purpose is to verify whether the objectives of the curriculum have been achieved. This is usually verified through the application of the experimental method. Experimentation, as described earlier, consists of designing and performing laboratory experiments on which the curriculum variable is manipulated and extraneous variables are controlled. The artistic mode, however, emphasizes process evaluation. It employs the technique of criticism for describing and interpreting the processes of the classroom. The goal of criticism is to render an evaluative judgment about the quality of classroom experience.

The third difference relates to the sources of evaluative data. Within the scientific approach, evaluative data are principally collected through the use of standardized instruments, particularly with norm-referenced achievement tests. In the artistic mode, the human instrument (the critic) is the major source of data. He/she is a connoisseur of classrooms who relies upon his/her perceptabilities, sensibilities, and knowledge gained through training and experience. According to Eisner (1985a), there are three advantages of using the human as instrument. First, many events that might be significant for those involved in the situation might not find a place on a standardized instrument. The human instrument pays attention to significant

events. Second, the human instrument is the only that can place events in a historical context. And third, the human instrument is able to recognize what is subtle, and yet significant. No other instrument, Eisner maintains, can do this.

The fourth difference is the form of representation used. The scientific mode employs a conventional or discursive form of representation, i.e. numbers. The data collected are expressed in quantitative terms. Quantification, according to Gay (1980), provides more precise and unambiguous descriptions than the verbal statement. He also claims that quantitative statements offer much more information, and provide an objective basis for comparison. Artistic approaches to evaluation rely upon the use of nondiscursive forms of representation; e.g., poetic language. According to Eisner (1985a), nondiscursive forms inform not by pointing but through intimation. They render an account of an experience which generates excitement.

The fifth dimension on which these approaches to curriculum evaluation differ is in their ultimate aim. The scientific mode aims at producing nomological generalizations (laws) that can provide dependable basis for predicting and controlling the outcomes of the curriculum. Scientific generalizations, as Koetting (1984) notes, are considered to be "truth statements outside of time and specific context" (p. 9). According to Eisner (1983), the formulation of laws for prediction and control has been a

dream of educational scientists since the early 1900s. He then argues that the current over-concern for efficiency and effectiveness in the school is nothing else than a reflection of such a dream. The artistically-oriented approach to evaluation does not aim to control or to produce predictions. It is after understanding. It is believed that with such understanding the ability of individuals to grasp and deal with situations like those portrayed in the evaluation will be increased (Eisner, 1985a).

The last difference between these two approaches lies in the criteria used for judging the validity of the evaluation. The scientific mode is concerned with the question of objectivity. To meet the criterion of objectivity is a requisite for establishing trust in the evaluative findings. An objective evaluation is defined as one which yields findings that are a function solely of the conditions of the evaluation and not of the interests, motives, and values of the evaluator (Guba and Lincoln, 1981). The artistic approach to evaluation, on the contrary, acknowledges the influence of the evaluator. The artistic evaluator, according to Eisner (1985b), is neither a neutral observer nor a disinterested interpreter. An artistic evaluation receives validation, then, through the processes of structural corroboration and referential adequacy. Structural corroboration establishes the degree of coherence of the evaluation. The evaluation meets this criterion of coherence if the pieces of evidence support one another and

make sense. In order for the evaluation to be referentially adequate, it must meet the criterion of usefulness. A useful evaluation is one which allows others to experience the situation portrayed in a new, more enlightened way.

These six differences between scientific and artistic approaches to curriculum evaluation provide support for this conclusion about their incompatibility. According to Guba (1983), in a competition between incompatible modes one is "dealing with an either-or proposition in which you 'puts yer money and yer takes yer cherce!'" (p.3).

Recommendations for Future Studies

As indicated in the introduction of this study, there are four competing paradigms in the field of curriculum evaluation, namely, the dominant natural science paradigm, and the legal, the anthropological, and the aesthetic emerging alternatives. One significant contribution of the present study is that, by identifying this competition, it provides the rationale for conducting future studies of paradigms in curriculum evaluation.

This study has already examined the philosophical, conceptual, and methodological components of the natural science and the aesthetic paradigms. Other paradigmatic components need to be examined. For instance, future studies might emphasize the psychological and sociological components. Other studies might explore the ideologies behind these paradigms. In deciding the focus for

ideological studies particularly, Michael Apple's suggestions might provide some guidance. For example, in his paper "Commonsense Categories and Curriculum Thought" (1975a), he suggests that the language categories used by positivists serve to legitimate the dominant values of the economic system. Related to the aesthetic perspective, Apple (1978) argues, with the cultural critic George Lukacs, that the "true bearers of ideology in art are the very forms, rather than the abstractable content, of the work itself" p. 506). He asks, for instance, why curriculum critics have chosen specifically literary forms for curriculum evaluation? He insists that these forms need to be ideologically analyzed. They too may serve important economic functions.

It is also recommended that studies be undertaken directed to explore how interrelated fields, e.g., curriculum development, instruction, and supervision, will be affected by accepting the aesthetic paradigm in the curriculum evaluation field. It would be significant to explore some of the problems of adopting this alternative paradigm. For instance, is it a political problem to bring this kind of change in evaluation, as Eisner (1985a) suggests? Is it an economic problem? Who shall make the decision of change? These areas deserve consideration.

The need for studying the other paradigms, i.e., the legal and the anthropological, has been already suggested. In particular, these studies could focus on the aspects rec-

ommended above in order to provide a common basis for comparison.

The primary intent of future studies should be to help curriculum evaluators make an informed choice of the paradigm that would guide the future of the field of curriculum evaluation. It is hoped that the present study of paradigms would make some contribution to that end.

A SELECTED BIBLIOGRAPHY

Anastasi, A. Psychological testing. New York: MacMillan Publishing Company, 1976.

Anderson, S. B., et al. Encyclopedia of educational evaluation. San Francisco, California: Jossey-Bass Publishers, 1977.

Apple, M. W. Commonsense categories and curriculum thought. In J.B. MacDonald and E. Zaret (Eds.), Schools in search of meaning. Washington, D.C.: Association for Supervision and Curriculum Development, 1975a.

Apple, M. W. Scientific interests and the nature of educational institutions. In W. Pinar (Ed.), Curriculum theorizing: The reconceptualists. Berkeley, California: McCutchan Publishing Corporation, 1975b.

Apple, M.W. Ideology and form in curriculum evaluation. In G. Willis (Ed.), Qualitative evaluation: Concepts and cases in curriculum criticism. Berkeley, California: McCutchan Publishing Corporation, 1978.

Atkin, M.J. Behavioral objectives in curriculum design: A cautionary note. The Science Teacher, 1968, 35, 27-30.

Ayer, A.J. The a priori. In L. W. Summer and L. Woods (Eds.), Necessary truth: A book of readings. New York: Random House, 1969.

Bahm, A.J. Axiology: The science of values. Albuquerque, New Mexico: A.J. Bahm Publisher, 1984.

Baker, R.L. Curriculum evaluation. Review of Educational Research, 1969, 39, 339-358.

Barbour, I. Paradigms in science and religion. In G. Gutting (Ed.), Paradigms and revolutions: Applications and appraisals of Thomas Kuhn's philosophy of science. Notre Dame, Indiana: University of Notre Dame Press, 1980.

Benton, T. Philosophical foundations of the three sociologies. London: Routledge and Kegan Paul Ltd., 1977.

Berk, R.A. Introduction. In R.A. Berk (Ed.), Educational evaluation methodology: The state of the art. Maryland: The Johns Hopkins University Press, 1981.

Berleant, A. The aesthetic field: A phenomenology of aesthetic experience. Springfield, Illinois: Charles C. Thomas Publisher, 1970.

Bernstein, R.J. The restructuring of social and positical theory. United States: University of Pennsylvania Press, 1978.

Best, J.W. Research in education. Englewood Cliffs, New Jersey: Prentice-Hall Company, 1970.

Bobbitt, F. General outlines and the curriculum for the individual. In H.L. Caswell and D.A. Campbell (Eds.), Readings in curriculum development. New York: American Book Company, 1935.

Bridges, J.H. (T.). A general view of positivism, by Auguste Comte. New York: Robert Speller and Sons, 1957.

Campbell, D.T. and Stanley, J.C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Charlton, D.G. Positivist thought in France. London: Oxford University Press, 1959.

Cronbach, L.J. Course improvement through evaluation. Teachers College Record, 1963, 64, 672-683.

Cronbach, L.J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.

Culbertson, J.A. Perspective: Antecedents of the theory movement. Educational Administration Quarterly, 1981, 17, 25-47.

Deutsch, E. On truth: An ontological theory. Honolulu: The University Press of Hawaii, 1979.

Dobson, R. and Dobson, J. The language of schooling. Washington, D.C.: University Press of america, Inc., 1981.

Dobson, R.L., Dobson, J.E. and Koetting J.R. Looking at, talking about, and living with children: Reflections on the process of schooling. Lanham, MD: University Press of America, Inc., 1985.

Donmoyer, R. The evaluator as artist. In H.A. Giroux, A.N. Penna, and W.F. Pinar (Eds.), Curriculum and instruction: Alternatives in education. Berkeley, California: McCutchan Publishing Corporation, 1981.

DuBois, P.H. A history of psychological testing. Boston: Allyn and Bacon, 1970.

Durant, W. The story of philosophy. New York: Washington Square Press, 1961.

Effrat, A. Power to the paradigms. An editorial introduction. Sociological Inquiry, 1972, 42, 3-34.

Eisner, E.W. Educational objectives: Help or hindrance? School Review, 1967, 75, 250-266.

Eisner, E.W. Emerging models for educational evaluation, School Review, 1972, 80, 573-590.

Eisner, E.W. Educational connoisseurship and criticism: Their form and functions in educational evaluation. In G.F. Madaus, M. Scriven, and D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation. Higham, MA: Kluwer Academic Publishers, 1983.

Eisner, E.W. The art of educational evaluation: A personal view. Philadelphia, PA: The Falmer Press, 1985a.

Eisner, E.W. The educational imagination: On the design and evaluation of school programs. 2nd Ed. New York: Macmillan, 1985b.

Erwin, E. The concept of meaninglessness. Baltimore, Maryland: The Johns Hopkins Press, 1970.

Evans, J.W. Head Start: Comments on the criticisms. In F.G. Caro (Ed.), Readings in evaluation research. New York: Russell Sage Foundations, 1971.

Frondizi, R. What is value? An introduction to axiology. Illinois: Open Court Publishing Company, Inc., 1963.

Gay, L.R. Educational evaluation and measurement: Competencies for analysis and application. Columbus, Ohio: Charles E. Merrill Publishing, C., 1980.

Gay, L.R. Educational research: Competencies for analysis and applications. Columbus, Ohio: Charles E. Merrill Publishing, Co., 1981.

Gephart, W.J. Editorial. The CEDR Quarterly, 1976, 9, 2.

Giroux, H.A. Toward a new sociology of curriculum. In H.A. Giroux, A.N. Penna, and W.F. Pinar, (Eds.), Curriculum and instruction: Alternatives in education. Berkeley, California: McCutchan Publishing Corporation, 1981.

Greene, M. Landscapes of learning. New York: Teachers College Press, 1978.

Guba, E.G. The failure of educational evaluation. Educational Technology, 1969, 9, 29-38.

Guba, E.G. Toward a methodology of naturalistic inquiry. Monograph Series, No. 8. Los Angeles, California: Center for the Study of Evaluation, University of California, 1978.

Guba, E.G. The context of emergent paradigm research. Paper presented at a Career Development Seminar sponsored by the Center for Public Affairs and the School of Education of the University of Kansas, and the University Council for Educational Administration, Overland Park, Kansas, November, 1983.

Guba, E.G. and Lincoln, Y.S. Effective evaluation. San Francisco, California: Jossey-Bass Publishers, 1981.

Guba, E.G. and Lincoln, Y.S. Epistemological and methodological bases of naturalistic inquiry. In G.F. Madaus, M. Scriven, and D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation. Higham, MA: Kluwer Academic Publishers, 1983.

Habermas, J. Knowledge and human interests. Boston: Beacon Press, 1971.

Hammond, R.L. Evaluation at the local level. In B.R. Worthen and J.R. Sanders (Eds.), Educational evaluation: Theory and practice. Belmont, California: Wadsworth Publishing Company, Inc., 1973.

Hartnack, J. Wittgenstein and modern philosophy. New York: Doubleday & Company, Inc., 1965.

Hasenmueller, C. Panofsky, iconography, and semiotics. Journal of Aesthetics and Art Criticism, 1978, 36, 289-301.

Hilliard, A.G. Democracy in evaluation: The ovolution of an art-science in context. In P.L. Hosford (Ed.), Using what we know about teaching. Alexandria, Virginia: Association of supervision and Curriculum Development, 1984.

Hofstadter, A. Truth and art. New York: Columbia University Press, 1965.

House, E.R. Assumptions underlying evaluation models. Educational Researcher, 1978 7, 4-12.

Huebner, D. Curriculum as a field of study. In H.F. Robinson (Ed.), Precedents and promise in the curriculum field. New York: Teachers College Press, 1966.

Huebner, D. Curricular language and classroom meanings. In W. Pinar (Ed.), Curriculum theorizing: The reconceptualists. Berkeley, California: McCutchan Publishing Corporation, 1975.

Hyman, J.L. and Napier, R.N. Evaluation in the schools: A human process for renewal. Monterey, California: Books/Cole Publishing Co., 1975.

Isenberg, A. The problem of belief. Journal of Aesthetics and Art Criticism, 1955, 13, 395-407.

Jackson, P. Curriculum and its discontents. Curriculum Inquiry, 1980, 10, 159-172.

Jenkins, D. and O'Toole, B. Curriculum evaluation, literary criticism, and the paracurriculum. In G. willis (Ed.), Qualitative evaluation; Concepts and cases in curriculum criticism. Berkeley, California: McCutchan Publishing Corporation, 1978.

Johnson, M. Definitions and models in curriculum theory. Educational Theory, 1967, 17, 127-140.

Johnson, M. On the meaning of curriculum design. In J.R. Gress and D.E. Purpel (Eds.), Curriculum: An introduction to the field. Berkeley, California: McCutchan Publishing Corporation, 1978.

Jones, K. Some epistemological considerations of paradigm shifts: Basic steps toward a formulated model of alternation. Sociological Review, 1977, 25, 253-271.

Kelly, E.F. Curriculum criticism and literary criticism: Comments on the analogy. In G. Willis (Ed.), Qualitative evaluation: Concepts and cases in curriculum criticism. Berkeley, California: McCutchan Publishing Corporation, 1978.

Kliebard, H.M. The Tyler rationale. In J.R. Gress and D.E. Purpel (Eds.), Curriculum: An introduction to the field. Berkeley, California: McCutchan Publishing Corporation, 1978.

Kneller, G.F. Introduction to the philosophy of education. New York: John Wiley & Sons, 1984.

Koetting, J.R. Foundations of naturalistic inquiry: Developing a theory base for understanding individual interpretations of reality. Media and Adult Learning, 1984, 6, 8-18.

Kraft, V. The Vienna circle: The origins of neo-positivism. New York: Philosophical Library, 1953.

Kuhn, T.S. The structure of scientific revolutions. 2nd Ed. Chicago: The University of Chicago Press, 1970.

Lavine, T.Z. From Socrates to Sartre: The philosophic quest. New York: Bantam books, Inc., 1984.

LeCompte, M.D. and Goetz, J.P. Problems of reliability and validity of ethnographic research. Review of Educational Research, 1982, 52, 31-60.

Levine, M. Scientific method and the adversary model: Some preliminary thoughts. American Psychologist, 1974, 29, 661-667.

Lewy, A. The practice of curriculum evaluation. Curriculum Theory Network, 1973, 11, 6-33.

Lewy, A. Handbook of curriculum evaluation. New York: Longman, Inc., 1977.

MacDonald, J.B. Educational models for instruction. In J.B. MacDonald and R.R. Leeper (Eds.), Theories of instruction. Washington D.C.: Association for Supervision and Curriculum Development, 1965.

MacDonald, J.B. and Wolfson, B.J. A case against behavioral objectives. The Elementary School Journal, 1970, 71, 119-128.

Madaus, G.F., Airasian, P.W. and Kellagham, T. School effectiveness: A reassessment of the evidence. New York: McGraw Hill Book Company, 1980.

Madaus, G.F., Stufflebeam, D., and Scriven, M.S. Program evaluation: A historical overview. In G.F. Madaus, M. Scriven, and D.L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation. Higham, MA: Kluwer Academic Publishers, 1983.

Mager, R. Preparing instructional objectives. California: Fearon Publishers, 1962.

Mann, J.S. Curriculum criticism. Teachers College Record, 1969, 71, 27-40.

Mazza, K. Reconceptual inquiry as an alternative mode of curriculum theory and practice: A critical study. Journal of Curriculum Theorizing, 1982, 4, 5-89.

McCutcheon, G. On the interpretation of classroom observations. Educational Researcher, 1981, 10, 5-10.

McLaughlin, M.W. Evaluation and reform. Cambridge, Mass.: Ballinger Publishing Company, 1975.

Metfessel, N.S. and Michael, W.B. A paradigm involving multiple-criterion measures for the evaluation of the effectiveness of school programs. Educational and Psychological Measurement, 1967, 27, 931-943.

Mill, J.S. A system of logic. New York: Harper and Row Publishers, 1873.

Morris, V.C. Philosophy and the American school. Boston: Houghton Mifflin Company, 1961.

Nitko, A.J. Defining "criterion-referenced test." In R.A. Berk (Ed), A guide to criterion-referenced test construction. Baltimore, Maryland: The Johns Hopkins University Press, 1984.

Owens, T.R. Educational evaluation by adversary proceeding. In E.R. House (Ed.), School evaluation: The politics and process. Berkeley, California: McCutchan Publishing Corporation, 1973.

Pap, A. The linguistic theory of logical necessity. In L.W. Summer and J. Woods (Eds.), Necessary truth: A book of readings. New York: Random House, 1969.

Parlett, M. and Hamilton, D. Evaluation as illumination: A new approach to the study of innovatory programmes. In D. Hamilton et al. (Eds.), Beyond the numbers game: A reader in educational evaluation. Berkeley, California: McCutchan, 1977.

Patton, M. Q. Alternative evaluation research paradigm. Grand Forks, North Dakota: North Dakota Study Group on Evaluation, 1975.

Patton, M. Q. Utilization-focused evaluation. Beverly Hills, California: Sage Publications, Inc., 1978.

Patton, M. Q. Qualitative evaluation methods. Beverly Hills, California: Sage Publications, Inc., 1980.

Pepper, S. C. The basis of criticism in the arts. Cambridge, Massachusetts: Harvard University Press, 1945.

Pinar, W. F. The reconceptualization of curriculum studies. Journal of Curriculum Studies, 1978, 10, 205-214.

Popham, W. J. Educational evaluation. Englewood Cliffs, New Jersey: Prentice-Hall, 1975.

Popkin, R. H. and Stroll, A. Philosophy made simple. New York: Doubleday and Company, Inc., 1956.

Popper, K. R. The logic of scientific discovery. New York: Science Editions, Inc., 1961.

Provus, M. Discrepancy evaluation. Berkeley, California: McCutchan Publishing Company, 1972.

Quine, W. V. Two dogmas of empiricism. In L. W. Summer and J. Woods (Eds.), Necessary truth: A book of readings. New York: Random House, 1969.

Reichardt, C. S. and Cook, T. D. Beyond qualitative and quantitative methods. In T. D. cook and C. S. Reichardt (Eds.), Qualitative and quantitative methods in evaluation research. Beverly Hills, California: Sage Publications, Inc., 1979.

Reid, L. A. Art, truth, and reality. The British Journal of Aesthetics, 1964, 4, 321-331.

Rice, J. M. The futility of the spelling grind. The Forum, 1897, 23, 163-172.

Riestra, M. A. Fundamentos filosoficos de la educacion. Rio Piedras, Puerto Rico: Editorial Universitaria, 1970.

Rogers, V. R. Qualitative and aesthetic views of curriculum and curriculum making. In A. Molnar (Ed.), current thought on curriculum. Alexandria, Virginia: Association for Supervision and Curriculum development, 1985.

Rossi, P. H. Standards for evaluation practice. San Francisco, California: Jossey-Bass, Inc., 1982.

Russell, B. The problems of philosophy. London: Oxford University Press, 1959.

Schroyer, T. The critique of domination: The origins and development of critical theory. Boston: Beacon Press Books, 1973.

Scriven, M. The methodology of evaluation. AERA Monograph Series on Curriculum evaluation, No. 1. Chicago: Rand McNally, 1967.

Scriven, M. Prose and cons about goal-free evaluation. In W. J. Popham (Ed.), Evaluation in education: Current applications. Berkeley, California: McCutchan Publishing Company, 1974.

Sequel, M. L. The curriculum field: Its formative years. New York: Teachers College, Columbus University Press, 1966.

Smith, J. K. Quantitative versus qualitative research: An attempt to clarify the issue. Educational Researcher, 1983, 12, 6-13.

Stake, R. E. The countenance of educational evaluation. Teachers College Record, 1967, 68, 523-540.

Stake, R. E. Program evaluation, particularly responsive evaluation. In G. F. Madaus, M. Scriven, and d. L. Stufflebeam (Eds.), Evaluation models: Viewpoints on educational and human services evaluation. Higham, MA: Kluwer Academic Publishers, 1983.

Stufflebeam, D. L. An introduction to the PDK book. In B. R. Worthen and J. R. Sanders (Eds.), Educational evaluation: Theory and practice. Belmont, California: Wadsworth Publishing Company, Inc., 1973.

Stufflebeam, D. L. et al. Educational evaluation and decision making. Itaska, Illinois: Peacock Publishers, Inc., 1971.

Stufflebeam, D. L. and Madaus, G. F. The standards for evaluation of educational programs, projects, and materials: A description and summary. In G. F. Madaus, M. Scriven, and D. L. Stufflebeam (Eds.), Evaluation models: Viewpoints one educational and human services evaluation. Higham, MA: Kluwer Academic Publishers, 1983.

Stufflebeam, D. L. and Webster, W. J. An analysis of alternative approaches to evaluation. Educational Evaluation and Policy Analysis, 1980, 2, 15-19.

Tranel, D. D. A lesson from the physicists. Personnel and Guidance Journal, 1981, 59, 425-429.

Tyler, R. W. General statement on evaluation. Journal of Educational Research, 1942, 35, 492-501.

Tyler, R. W. Basic principles of curriculum and instruction. Chicago: University of Chicago Press, 1949.

Unruh, G. G. and Unruh, A. Curriculum development: Problems, processes, and progress. Berkeley, California: McCutchan Publishing Corporation, 1984.

Vallance, E. Scanning horizons and looking at weeds: A critical description of the 'The great plains experience.' In G. Willis (Ed.), Qualitative evaluation: Concepts and cases in curriculum criticism. Berkeley, California: McCutchan Publishing Corporation, 1978.

Vargas, J. S. Writing worthwhile behavioral objectives. New York: Harper and Row, 1972.

Von Wright, G. H. Explanation and understanding. London: Routledge and Kenan Paul, 1971.

Webster's New Collegiate Dictionary. Springfield, Massachusetts: G. & C. Merrian Company, 1979.

Westbury, I. Curriculum evaluation. Review of Educational Research, 1969, 40, 239-257.

White, A. R. Truth. New York: Doubleday & Company, Inc., 1970.

White, M. The age of analysis: 20th century philosophers. New York: New American Library, 1955.

Willis, G. Curriculum criticism and literary criticism. In G. Willis (Eds.), Qualitative evaluation: Concepts and cases in curriculum criticism, Berkeley, California: McCutchan Publishing Corporation, 1978a.

Willis, G. Qualitative evaluation as the aesthetic, personal, and political dimensions of curriculum criticism. In G. Willis (Ed.), Qualitative evaluation: Concepts and cases in curriculum criticism. Berkeley, California: McCutchan Publishing Corporation, 1978b.

Willis, G. and Allen, A. J. Patterns of phenomenological response to curricula: Implications. In G. Willis (Ed.), Qualitative evaluation: Concepts and cases in curriculum criticism. Berkeley, California: McCutchan Publishing Corporation, 1978.

Wilson, S. The use of ethnographic techniques in educational research. Review of Educational Research, 1977, 47, 245-265.

Wolf, R. L. The application of select legal concepts to educational evaluation. (Unpub. Ph.D. dissertation, University of Illinois, 1974.)

Wolf, R. L. Trial by jury: A new evaluation method. Phi Delta Kappan, 1975, 57, 185-197.

Worthen, B. R. and Sanders, J. R. Educational evaluation: Theory and practice. Belmont, California: Wadsworth Publishing Company, Inc., 1973.

Worthen, B. R. Characteristics of good evaluation studies. Journal of Research and Development in Education, 1977, 10, 3-20.

Worthen, B. R. and Rogers, W. T. Pitfalls and potential of adversary evaluation. Educational Leadership, 1980, 37, 536-543.

Zais, R. S. Curriculum: Principles and foundations. New York: Harper and Row Publishers, 1976.

VITA

Gladys E. Pereles

Candidate for the Degree of

Doctor of Education

Thesis: AN EXAMINATION OF THE DOMINANT AND THE ALTERNATIVE PARADIGMS IN CURRICULUM EVALUATION

Major Field: Curriculum and Instruction

Biographical:

Personal Data: Born in Ponce, Puerto Rico, January 25, 1946, the daughter of Rosa Velez and Genaro Pereles.

Education: Graduated from Ponce High School, Ponce, Puerto Rico, in May, 1964; received Bachelor of Science degree in Secondary Education from Catholic University of Puerto Rico in May, 1972; received Master of Education degree in Administration and Supervision from Catholic University of Puerto Rico in May, 1976; completed requirements for the Doctor of Education degree at Oklahoma State University in May, 1987.

Professional Experience: Internship in Science of Secondary School, Dr. Pila High School, Ponca, Puerto Rico, 1971-1972; Internship in Administration and Supervision, InterAmerican Regional College, Ponce, Puerto Rico, 1975-1976. Seventh, eighth, and ninth grade science and math teacher at Pujals Junior High School, Ponce, Puerto Rico, 1972-1978; eleventh grade chemistry teacher at Armstrong High School, Ponce, Puerto Rico, 1979-1981. Laboratory Technician I, Physiological Sciences, Oklahoma State University, Stillwater, Oklahoma, 1981-1982; Test Scorer, Bureau of Tests and Measurements, Oklahoma State University, Stillwater, Oklahoma, 1982-1984; Student Technician, Microbiology Department, Oklahoma State University, Stillwater, Oklahoma, 1985.