

A CRITICAL ANALYSIS OF
THE PORTFOLIO METHOD
AS AN ALTERNATIVE
ASSESSMENT FORM

By

SANDRA DOUGLAS MASSEY

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

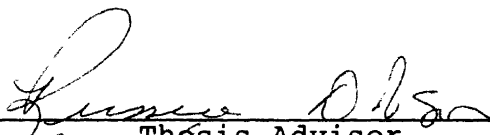
1969

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 1993

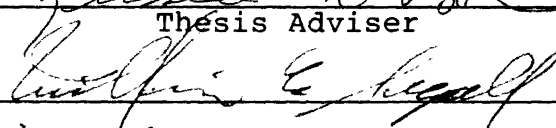
OKLAHOMA STATE UNIVERSITY

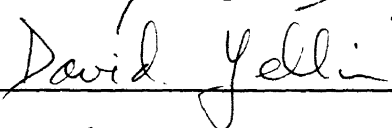
A CRITICAL ANALYSIS OF
THE PORTFOLIO METHOD
AS AN ALTERNATIVE
ASSESSMENT FORM

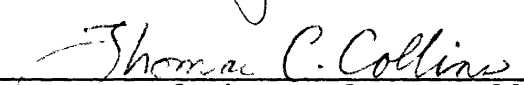
Thesis Approved:



Thesis Adviser







Dean of the Graduate College

ACKNOWLEDGMENTS

This thesis is the result of the efforts of many people. First and foremost, I would like to express my sincere appreciation to Dr. Russell Dobson for his encouragement, advice and mentorship throughout this project. My thanks and appreciation also go to Dr. David Yellin and Dr. William Segall for serving on my committee and for their suggestions, support and patience.

To Dr. Randall Koetting I extend my gratitude for introducing me to the field of curriculum studies and for "kindling the fire." To Dr. Larry Foreman goes my thanks for helping me talk through the initial stages of this paper. To the faculty of Oklahoma State University I extend my thanks for informative and stimulating classes that gave me an opportunity to get an education - not just a degree.

On a more personal note, I would like to thank the following people: my parents, Maxine and Pete Douglas, who taught me the value of education and helped me develop a love of learning; my children, Jennifer and Scott, who shared our time together so I could pursue my dream; and most of all to my husband, Lee, who provided moral support and infinite patience throughout this degree program. Without his love and encouragement, as my best friend, this endeavor would not have been the success that it has been.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.	1
The Role of Assessment	1
The Value of Orientation Assessment.	3
The Field of Assessment and Evaluation	4
Quantitative and Qualitative Assessment.	5
Organization of the Study.	8
Limitations, Method and Purpose.	11
II. FOUNDATIONS OF THE DOMINANT PARADIGM.	13
The Language of Control.	13
Historical Roots	14
The Greek Influence	14
The Roman Influence	16
The Medieval Influence.	17
The Renaissance Influence	18
The Protestant Reformation Influence.	19
Democratic and Religious Roots	20
The Colonial Religious Influence.	21
The Social and Political Influence.	22
Philosophical Roots.	25
Traditional and Contemporary Philosophies:	
Purposes and Roles	27
Traditional Philosophies.	27
Contemporary Philosophies	31
A View of Humanity	35
The Role of Schools.	37
The Overt and Covert Curriculum	38
Goals, Functions and Aims	39
Goals As a Socio-Political Process.	40
The Language and Metaphors of Schooling.	42
Metaphors As Shapers of Social Reality.	43
The Dominant Metaphor: The Factory Metaphor.	46
The Conflict of Articulation	48
III. CONCERNS, CRITICISMS, AND DEFENSES OF ASSESSMENT.	51
Assessment As Control.	51
Existing Assessment Practices.	53

Chapter	Page
Assessment as a Definition of Curriculum.	55
Desirable Characteristics of Assessment.	58
The Values Implied in Assessment	61
The Purposes of Assessment.	62
The Side-Effects of Assessment.	65
A Defensive Rejoinder Concerning Assessment.	73
Trivialization and Deskillling	79
An Aesthetic Alternative	80
IV. PORTFOLIOS AS AN ALTERNATIVE ASSESSMENT FORM. . .	86
Portfolios	87
Portfolios in the Classroom	89
Why Do We Need Portfolios?	89
What Is a Portfolio?	91
What Should Portfolios Contain?. .	92
How Much Should Portfolios Contain?	93
What Evidence Will the Portfolio Contain?	94
How Much Time Is Required for Portfolios?.	95
How and When Will the Classroom Teacher Assess a Student's Work in the Portfolio?	95
What Happens to the Portfolio at the End of a Term?	96
How Are the Grades and Evaluation Communicated?.	96
A Question of Grades.	97
Putting Portfolios in Perspective.	102
Test Content in Curriculum.	104
Judging the Quality of Assessment . . .	105
Other Concerns.	112
Assessment As An Occasion for Learning . . .	114
The Move Away from Standardized Tests	114
The Need for Empowerment of Teachers and Students.	115
The Recognition of Multiple Levels of Intelligence	116
The Need for a More Humanistic Language.	116
The Need for Self-Assessment.	117
The Need to Do Away with Labels	118
The Benefits of Portfolios	118
V. IMPLICATIONS, CONCLUSIONS AND RECOMMENDATIONS . .	122

Chapter	Page
Introduction	122
Implications	124
Conclusions.	128
Recommendations.	133
BIBLIOGRAPHY	138

CHAPTER I

INTRODUCTION

The Role of Assessment

Eliot W. Eisner, in The Art of Educational Evaluation: A Personal View (1985a), suggests that the most complex educational task of all is the "systematic and scientific development" of curriculum (Eisner, 1985). If that is true, then surely one of the most difficult aspects of that task is that of evaluation or assessment. There are many treatises, books, monographs, and so on written about educational philosophies and nearly all the writers address the issue of evaluation at one time or another.

In discussing the role of evaluation Ralph Tyler (1949) noted, evaluation

becomes a process for finding out how far the learning experiences as developed and organized are actually producing the desired results and the process of evaluation will involve indentifying the strengths and weaknesses of the plans...[evaluation will also] check the validity of the hypotheses the program is founded on...check the effectiveness of the instruments...note how the curriculum is effective and where it needs improvement." (p. 105)

Jerrold Kemp discusses the evaluation of learning as the last of his four essential elements in The Instructional Design Process (1985, p. 160). Kemp also points out that an important principle that contributes to successful learning is to provide feedback to the learner on how well they are learning, a process which includes student self-evaluation (p.178,179).

When many in America are calling for a change in schools and curriculum, "an essential element in the redesigning of schools to match changing world conditions is the way we assess student learning" (Wolf, Mahieu and Eresh; 1992, p. 7). This idea is not new. Even in 1949 Tyler observed that often people think of evaluation as synonymous with giving of paper and pencil tests (p.107). Eisner noted in 1985 that while there was significant concern with the so-called basics and their assessment, there was another movement developing that was concerned with the creation of a fundamentally different conception of education, in particular educational evaluation.

This movement was in direct conflict with the idea of education limited to the three R's and to the form of evaluation limited to quantitative description. Since evaluation involves getting evidence about behavior changes in students or objectives reached by students, then other forms of information about the desired behaviors and objectives provide valid and appropriate methods of evaluation, according to Tyler. While noting that students

are not asked all the questions possible about a principle or concept, Tyler pointed out that students are usually questioned on a sample of things and then their reaction to the total set of items that might be involved in their knowledge is inferred from their answers. He also noted that "it is assumed that it is possible to infer the person's characteristic performance by appraising his reaction in a sample of situations..." (p. 109). Eisner (1985a) agrees with Tyler's criticism by way of an interesting analogy which points out that:

To use such devices as the exclusive tools for evaluation is like casting a net into the sea that is intentionally designed to let the most interesting fish get away. To describe the ones that are caught strictly in terms of their weight and length is to reduce radically what can be known about them. To proceed further to conclude that the content of the sea consists of fish like those that remained in the net is to compound the error even further. (p.176)

The Value Orientation of Assessment

In fact, one of the first problems encountered with evaluation is in defining the difference, if any, between evaluation and assessment. In the minds of most, these two activities are synonymous. However, Eisner points out that evaluation deals with appraising the value of some object and is, without question, value oriented, while he implies

that the word assessment carries a more objective connotation.

Although the very nature of the word "value" makes it difficult to generalize a good definition, Selakovich (1967) cites Kluckhohn who provided a working definition in the statement: "A value is a selective orientation toward experience, implying deep commitment or repudiation, which influences the ordering of choices between possible alternatives in action." (p. 5)

One of the major problems that results from the value orientation aspect of Eisner's definition of evaluation is that, if it is true, educators, and specifically curriculum builders, are required to decide what is valuable and what is not. This then, becomes part of the ongoing debate concerning what schools "ought" to teach and ultimately how to evaluate what is taught.

The Field of Assessment and Evaluation

Schubert (1986) offers the observation that the category of evaluation has "emerged into a field of its own in the twentieth century" (p. 261). David Satterly (1989) defines educational assessment as "all processes and products which describe the nature and extent of children's learning; its degree of correspondence with the aims and objectives of teaching and its relationship with the environments which are designed to facilitate learning" (p. 3).

There are several philosophies that influence the assessment of individuals. At one end of the continuum are those who favor complete objective evaluation based on test scores and statistical averages, while at the other extreme are those who call for an end to testing and all forms of ranking and categorizing which might be used to classify and/or label an individual. Most who discuss the evaluation process do suggest that there is no one best method but most also acknowledge that the heavy reliance on standardized tests is, in a sense, limiting what is done in the classroom because of the tendency to "teach to the test". Maddeus (cited in Sears and Marshall, 1990) says "curriculum is narrowed to preparing for exams...test results should be only one element to be considered in deciding on curriculum goals ... the tendency to standardize the curriculum through the use of easy-to-administer tests should be resisted, even by the measurement experts." (p. 202)

Quantitative and Qualitative Assessment

With this in mind, this paper will seek to analyze the accusations and defenses leveled at standardized and normative testing in an attempt to see if the objectives claimed by the educational community are actually being met through the current style of evaluation common to most programs and curriculum. As Eisner (1985a) points out:

Evaluation has...been regarded as analogous to testing.

It has defined educational priorities in the curriculum

by virtue of the public status test scores have received. It has limited our understanding of the processes of educational practice by its neglect of the conditions that account for the outcomes that have been measured. It has neglected large areas of important educational outcomes by employing forms of representation that cannot describe certain significant features of a student's work. And it has encouraged teachers to focus upon teaching bits and pieces of information because of the ways in which tests have been constructed. (p.4)

Eisner also points out that "in the process [of giving the public what it seeks in terms of accountability] the children may be sacrificed educationally for the seductive comforts of high test scores" (1985a, p. 141). He also notes that one of the disappointing things about current federal efforts to reform the schools is the "apparent failure to use serious scholarship to cope with an admittedly complicated and seemingly intractable problem. It would be better to say 'we don't know' than to try to find the silver bullet that has so many times failed to hit its target" (Eisner, 1992, p. 723). Eisner also suggests along with new forms of assessment that those who use more qualitative methods do not reject quantitative procedures...what they should reject is the assumption that objectivity can only be secured through quantitative or scientific methods. He rejects the claim "implicit or

explicit- that rigour in educational inquiry requires the use of methods that result in conclusions that can be stated in terms of probabilities!" (1985, p. 136).

In cautioning care concerning the adoption of new methods before they, themselves, have been rigorously assessed, Richard Stiggins, director of the Center for Classroom Assessment at the Northwest Regional Educational Laboratory says:

Moving slowly may ultimately be a good thing (because some of the new tests may be oversold) ... one of the things that troubles me greatly is that we're setting up performance assessments and paper-and-pencil tests against one another ... each test has a contribution to make. We can't throw away any of the tools at our disposal (cited in O'Neil, 1992a, p.19).

On a more critical note, Michael Apple (1990) argues that "social and economic values are...[found] in the 'formal corpus of school knowledge' we preserve in our curricula, in our modes of teaching, and in our principles, standards, and forms of evaluation" (p. 9).

Assessment and evaluation is not limited to only students in the classroom, but also to the program, the school and the teachers. In Assessment in Schools, Satterly (1989) points out that "If one's concept of effective teaching incorporates adaptation to individual children based on their strengths and weaknesses rather than the treatment of the class as a unit, some form of assessment is

required...it also involves assessment of objectives and strategies of the teacher" (p. 6).

The problem facing the educational world in general and the curriculum specialists in particular is that of deciding what form evaluation should take. As Sund and Trowbridge (1974) note:

Probably nothing is so well known and so little understood by the teachers as evaluation. Evaluation involves the total assessment by the instructor of a student's learning and development including understanding of cognitive critical thinking processes, subject matter, competence, multiple talents, values, self-concept, laboratory skills, and the ability and willingness to work. (p.242)

There are a number of assessment alternatives available ranging from outcomes, grades, and standardized tests to informal and formal assessments, behavioral objectives, goal free evaluation and others. This paper will be concerned with the use of one form of assessment currently receiving some notoriety - the portfolio.

Organization of the Study

The chapters that follow will examine three major areas of inquiry concerning evaluation and assessment; the historical background and the way society views humans, assessment as it now stands (based on grades and standardized tests) and portfolios as an alternative form of

assessment that might meet the stated objectives of all concerned in the education process.

Chapter Two will look at the historical and philosophical foundations which have led to the model of assessment currently in use. This chapter will also focus on how this nation views the school's role in society - exactly what it is that people want schools to do and how do they think schools should meet these expectations? Although a number of educational theories exist today, some major distinctions as defined by Dobson, Dobson and Koetting (1985) will serve as a basis for comparison and contrast.

Chapter Two will also examine the dominant metaphor used to describe education methods and the philosophies out of which it has developed. Assessment procedures generally find their methods and formats in the prevailing educational philosophy of the times. But as Michael Apple (1990) has said "the kinds of values and rules that educators use to evaluate their students' "success" and "failure" determines their own ideological position and the functioning of their theories, principles, and modes of organization" (p. 111). Therefore, according to Apple, curriculum design, the creating of educative environments in which students are to dwell, is inherently a political and moral process (Apple, 1990, p.111). It is to the literature of these varying philosophies that Chapter Two will be addressed.

Chapter Three will discuss desirable characteristics of evaluation and examine the existing assessment practices.

This chapter will also look at the advocacy and criticism aimed at standardized and norm-referenced testing. In addition, this chapter will consider an aesthetic alternative as proposed by Elliot Eisner.

Chapter Four will look in detail at the portfolio approach to assessment and the philosophies upon which this approach is based. Since there are so many variations of philosophy concerning evaluation it would be prohibitive to cover all of them in this single treatment. Those observations made by Eisner, Goodlad, Gardner, Wiggins and Apple will form the basic structure for the discussion along with ideas from other authors who support these viewpoints.

Some basic conclusions involving the implications of changing how the American education system looks at evaluation and some suggestions about future studies to determine the value of this evaluation method will be the focus of Chapter Five. Since there are currently few completed long-range studies available upon which to base an all-inclusive evaluation policy, the need is obviously there to conduct some pilot projects using a combination of the best methods available for study at this time. While the intent is to find the best method of evaluation, the caution is to go slowly and act rather than react. Often the way that education seems to deal with problems is to throw everything out and begin again. There is the tendency to want so much to do what is right that sweeping changes are made without giving thought to keeping what works.

Educators are constantly "reinventing the wheel" in an attempt to show the public that education is not stagnate, nor oblivious to the needs of its constituents.

Limitations, Method and Purpose

This study is limited in that it will not conclude with a definitive answer to the problems of education. The intent is to look at the ideological foundations on which educators, according to Dobson, Dobson and Koeting (1985), ordinarily "look at, [think about], and talk about" children in particular and education in general. The method of this study is to review these foundations, analyze their influence and discuss their implications.

The purpose of this paper is to provide a critical analysis which looks at the historical perspectives that have influenced how we evaluate students and then looks at the portfolio as one alternative for assessment and evaluation. While it certainly offers suggestions to the reader, it is written more to define a direction for the writer. It is a process of "discovering, not having discovered" as noted by Schubert (1986). Final answers are in the process of being created and reconstructed to fit the needs of changing circumstances (Schubert, 1986, p.2). This study is an attempt to conduct a learning situation and to develop an understanding and consideration of assessment alternatives.

As Gerald Grant (1991) observes:

Only a century ago, to be literate was to sign your name or read a highly familiar text, but neither of these definitions is sufficient today. Similarly, if we are to ensure student learning, we will have to conduct ourselves as learners in developing alternatives to standardized testing. We have to push beyond generating engaging alternatives by listening to critiques and revising and improving our own portfolio of approaches. It will take what the painter, Ben Shaw, saw as the heart of good artistry: "The capacity to be the spontaneous imaginer and the inexorable critic" - not once, but iteratively - as our culture shifts and our understanding deepens. (p. 64)

CHAPTER II

FOUNDATIONS OF THE DOMINANT PARADIGM

The Language of Control

In order to understand dominant education paradigms and to understand the culture which influences the development of these paradigms, it is necessary to look at the predominant theories which have influenced education, the current models now in use and the language metaphors which govern these models. The current model, that of the factory, is one of control and management and is a result of the emphasis on scientific methods and the influence of the industrial model growing out of the management system of the Industrial Revolution.

In the vocabulary of this model is the word "control". As a verb, the word "control" means "to exercise restraint or direction over; dominate; command." In addition, "control" means "to test or verify by a parallel experiment or other standard of comparison." As a noun, the word refers to "the act or power of controlling; the situation of being controlled; a standard of comparison; a device for regulating and guiding a machine; the prevention of the flourishing or spread of something undesirable." Is it any

wonder that the predominant activity in most schools today according to Apple (1990), Goodlad (1984) and others is that of control. The question is: How did the educational community arrive at this stage in which the locus of control resides within the realm of the school? More importantly, the question is with whom in the school does the most control lie?

To arrive at an answer it is necessary to look at the early development of education in America, but the search should begin further back than America's colonial period. To truly understand the foundations of the American form of education one must look to the beginnings of the culture in Greece and Rome because it was there that many of the ideas that govern our current philosophies actually began. It is important to look briefly at the contributions of not only Socrates, Plato and Aristotle, but to also follow the influence of Quintillian through to the Middle Ages and to note the scientific and reasoning influence of the Renaissance and Reformation. It is only then, after viewing the historical roots of schooling, that one can focus on the development of American schooling.

Historical Roots

The Greek Influence

Socrates was the founder of the "inquiry" method of teaching. Ellis, Cogan and Howey (1991) note that:

Socrates engaged citizen and slave, young and old, in

dialogue about issues of the day. He probed and questioned to uncover truth. Socrates believed that education and society were closely bonded. If education is successful and produces good citizens, then society will be strong and good. However, the failure of education yields a failed society as well, a society that is weak and whose workings are undermined. (p. 66)

A major issue then with Socrates was to discover what type of education produced good citizens. Today, citizenship education is included in most statements of purpose.

Plato, as a student of Socrates, was an idealist who believed, in his early years, that the purpose of education, according to Ellis et al. (1991), was:

to mold people who would be capable and devoted to serving the state; if the people were properly 'molded,' the state could be a utopian one. The key was that the persons needed to be matched to their job in terms of their ability. Thus, some were destined to do physical labor while others would lead and govern. (p. 67)

As Plato grew older his thinking also changed somewhat and he developed a more realistic outlook about education. He later established his Academia, a school where young men came to study. It was here that he realized that education did not take place only within the walls of the Academia,

but that all the environmental stimuli that children encountered were a part of education.

Probably it was Aristotle who had the greatest influence on the development of our colonial educational thought. He developed, as a student of Plato, some very basic beliefs concerning schooling. First, according to Ellis et al. (1991, p. 67), he believed that "education was so central to the preservation of the state that only the state should undertake the education of its citizens." Secondly, "he believed that there should be a common core of knowledge for all, a basic education for citizenship with a curriculum that would include reading, writing, music, and physical training." This belief is one of the major premises of education today and can be found in the prescribed curricula and the standards for accreditation and certification that are a fundamental part of every public school system and most private school systems. "Aristotle's Lyceum laid the groundwork for the fields we know today...as the classical humanist tradition" and required in most colleges and universities as a "liberal Arts" education.

The Roman Influence

The Romans, of course, drew upon the Greeks for the development of their own systems of educating and schooling. The Greek pedagogues, used as mentors and used to accompany the Greek boys through their schooling experience, were brought to the Roman Republic, not as tutors within the

families as was the custom in Greece, but were instead enlisted as teachers in the schools. Ellis et al. (1991) observe that:

The Roman contribution to the process of education was the "grammar school" and the compendium of studies which set the pattern for what we, today term a liberal arts education. These studies included logic, literature, music, geometry, architecture, grammar, rhetoric, history, and astronomy - the subjects that continue to be the underlying basis for liberal education. (p. 68)

Quintillian was a contributor to Roman schooling. He believed that children should be educated in their early years and that they would best learn, as noted by Ellis et al. (1991, p. 69), "by praise and positive models. He bitterly opposed the harsh treatment used by some of his contemporaries. He strongly believed that the child's special needs and interests should be taken into account when prescribing the curriculum for each individual."

The Medieval Influence

After the Roman empire crumbled, the center of culture was established in Constantinople which became the "repository of Western classical learning that would be rediscovered hundreds of years later by the Crusaders from Europe." (Ellis et al., 1991, p. 69) The West gradually modified the social, political, and military system to

include "interlocking rights, privileges, services and mutual protection. This system [was] called feudalism" (p. 69).

During the period known as the Middle Ages or the medieval period, education was largely undertaken by monastery schools established by the church so that "literacy was almost exclusively the domain of the Church." Great universities were established that "laid the foundation of the intellectual rebirth which was about to take place" (Ellis et al., 1991, p. 70).

The Renaissance Influence

As the crusaders returned from the East, they brought with them knowledge of a people who were in possession of great learning and literature. This knowledge and the literature associated with it "revitalized schooling throughout Italy" and eventually influenced the Northern European scholars who came there to learn. "Religious reform" comment Ellis et al. (1991):

was frequently bound up with civil reform and with anti-church power disputes. As the pace of life quickened, the interaction of people and the collision of ideas brought out the complexities of issues that could not be resolved by simple authoritative statements. In emulation of classical practice, the Renaissance thinkers became more confident of humanity's ability to use reason to resolve problems.

(p. 71)

The Protestant Reformation Influence

In addition, the Protestant Reformation in the fourteenth, fifteenth and sixteenth centuries had a great impact upon education. Martin Luther's Ninety-Five Theses became important to the development of literacy since the premise was that people needed to be able to read and to interpret the Bible for themselves, rather than to simply accept the Church's position. As Ellis et al. (1991, p. 71) point out " the authority of the Roman Church was broken and the foundations for secularized education were established."

During this period in which science and reasoning became the means to understand man and his world, there were a number of notable influences on education. Ellis et al. (1991) have noted such contributors as:

Francis Bacon, the English philosopher who developed the method of scientific inquiry used today by many scholars; Johann Comenius, a Moravian bishop who advocated universal education for both boys and girls and active use of sensory stimuli in learning experiences; John Locke, the English philosopher whose tabula rosa theory was the basis for modern behavior psychology; Jean Jacques Rousseau, who contributed to the theory of developmental psychology through his novel Emile, which describes the developmental nature of children; John Pestalozzi, a Swiss educator who

tested Rousseau's ideas in his experimental schools through the use of 'object lessons' designed to develop principles of teaching in elementary school; Frederich Froebel, who was a follower of Pestalozzi and who established the first kindergarten in 1837 which emphasized activity-based curricula and teaching methods; Johann Herbart, a German philosopher who developed a psychological theory of learning which resulted in five formal steps of instruction (preparation, presentation, association, generalization and application); and Charles Darwin, whose theory of evolution ... is still a center of controversy today. (p. 74)

These persons, through their writings, were to have a profound impact on education and schooling in the colonies of the new world. "These antecedents are essential to understanding education programs and practices in the New World" (Ellis et al., 1991, p.74).

Democratic and Religious Roots

A free public education was not always the case in American life. The principle of public education was left to individual public demands and the needs of the society. "Yet this nation was founded on democratic principles; and a society imbued with such ideals of democracy could not long survive without an educational system designed to perpetuate a citizenry that would understand and uphold its ideals"

(Ellis et al., 1991, p. 78).

The Colonial Religious Influence

How did our current idea and method of schooling develop as it exists today? To begin to answer this question it is necessary to take a brief look at American education from the early colonial religious schools to the integrated public schools of today.

Just as religion was one of the primary motives for the emigration of people to what would become the United States of America, so too was it one of the primary motives for the creation of schools. The first colonial schools were church-related and the education in these schools was primarily religious in character and content. One of the first schooling laws to pass in an ever increasing endeavor to legislate educational policy and equality was the famous Massachusetts Bay Colony law of 1642 which mandated that parents were to ensure that their children could read to "understand religious principles and laws of the Colony" (Ellis et al., 1991, p. 78). This was followed in 1647 by the "Old Deluder Satan Law", which required that every community of 50 or more households must provide a teacher to instruct children in reading and writing.

There was still a distinction between the Latin grammar schools which provided for the education of the wealthy in a classical tradition and the common schools which provided for the education of the masses. These common schools

generally taught the basic subjects on an individual basis and most of the students, who went on to become tradesmen or clerks, seldom progressed beyond this elementary level of education. Secondary and higher education in the colonies were usually reserved for the wealthy class and were for the purpose of preparing boys for college or for the ministry. The colleges that were founded in the colonies used a classical approach to education and only the select (usually governed by wealth) ever were admitted. The common schools that existed for the benefit of the general public were paid for out of local taxes and thus began the concept and practice of public support for education.

When America sought independence from England, education and schooling in general were somewhat ignored in the endeavor engage every citizen in the war effort. Although the Constitution that was developed by this emerging nation did not specifically address educational issues, the Bill of Rights did specify that "all powers not delegated to the United States by the Constitution, nor prohibited by it to the states, are reserved to the states respectively, or to the People." Education in the United States today continues to be largely de-centralized which is a characteristic that distinguishes it from most educational systems in other parts of the world.

The Social and Political Influence

As America grew, American education continued to be an

area of increased focus and attention. The period of 1800-1865 was a period of interest in universal schooling. John Goodlad (1979) notes that the seventeenth century focused on narrowly academic and religious goals, while the eighteenth and nineteenth centuries added vocational and social goals. The period from 1865-1920 saw the end of the Civil War, the rapid industrialization of the North and massive immigration which brought people who needed to be "enculturated". Public schools were the answer in terms of efficiency and economy. As school grew it became necessary to look to quality assurance and to maintain standards across the nation and across all levels. This opened the way for the development of evaluation objectives, criteria and methods. "Accreditation of schools," according to Ellis et al. (1991) led to the licensing or certification of teachers by the individual states where teachers in training were to follow a prescribed course of study.

As America found it necessary to protect its ideals of democracy, the nation developed ordinances and legislation regarding education and today schooling and all of its attendant functions are a large part of the American society. Ellis et al. (1991) note that the educational institution is supported by state, local and federal governments as well as by private sources. Its intent is to provide equal opportunity for all Americans, regardless of race, ethnic origin, or sex. It is regarded as vital to the future of our nation. In short, education is an

inseparable, guaranteed part of American life.

The period of expansion from 1865-1920 also became a period of major educational thought with respect to philosophy, psychology, and methodology in schools.

"William T. Harris, Francis W. Parker, Emma Willard, William Janes, Mary McLeod Bethune, G. Stanley Hall, Prudence Crandall, Edward L. Thorndike, and George S. Counts were major figures. However, perhaps the most influential educational thinker was John Dewey. These people stimulated education in the United States with new ideas and innovative practices. They were instrumental in beginning to articulate a truly American philosophy of education and schooling" (Ellis et al., 1991, p.86).

Ellis et al. (1991) also mention that "the last sixty-five years of schooling history in the United States are characterized by four major developments:

the progressive education movement; the 1954 United States Supreme Court decision declaring the 'separate but equal' doctrine unconstitutional; the major role in school curriculum development assumed by the federal government as a result of the "space race"; and a call for a return to "excellence" in education in the 1980s and beyond. (p.89)

Although there is some question as to whether one can return to an "excellence" that was perhaps never abandoned, it is still appropriate to look at the philosophical foundations that have influenced educators in the past and

that continue to provide an impetus to contemporary educational thought. The paradigms that currently are reflected in American schooling can be traced to these major areas of philosophical thought. With this in mind, the next part of this chapter will look at five major schools of philosophy that have influenced educational thought in the United States. These are not studies in depth, but rather definitions designed to give a general overview of each philosophy. It is hoped that this overview will provide an understanding of the metaphors which they have produced.

Philosophical Roots

According to Ellis et al. (1991) there are five major philosophical schools of thought that include:

Idealism. Idealism is regarded as one of the oldest schools of philosophic thought. Plato, who is generally regarded as the father of idealism in the West, lived approximately 2,500 years ago; since then, the philosophy has been propounded in various forms by many others. Idealism emphasizes moral and spiritual reality as the primary source for an explanation of the universe. Truth and values are seen as absolute and universal. Knowledge is in the mind, and needs only to be brought to the conscious level through introspection. To know is to rethink the latent ideas which are already present in the mind.

Realism. Realism is another of the classical schools of

thought. Aristotle contributed a great deal to the development of this philosophy in Greece. The realist sees the world in material terms. The world of things, which exists independently of the mind, can be revealed to the mind through sensory experience and the use of reason. The realist views reality in terms of the world of nature. Everything is derived from nature and is subject to its laws. Realism suggests that life in its physical, mental, moral, and spiritual sense is attributable to and explicable by the ordinary operations of the natural world. Realism is more concerned with things as they are than with things as they should be.

Neo-Thomism. Neo-Thomism is often referred to as scholasticism. It was developed by Saint Thomas Aquinas in the mid-thirteenth century. This Christian philosopher integrated Christian thought with that of the early Greeks to try and bridge the gap between dualism, realism and idealism. Humanity was viewed as having both mind and body with man as the ultimate creation of God. Aquinas suggested that man needed both reason and faith to understand God and the universe. However, absolute truth was to be found in faith. This philosophy is still influential today as part of the official doctrine of the Roman Catholic Church.

Experimentalism/Pragmatism. This philosophy grew out of the work of the English philosopher Sir Francis Bacon and the German philosopher Immanuel Kant. This philosophy views reality as constantly changing; thus reality can only be known through experience. According to the pragmatist there is not absolute or permanent knowledge; only that which can be observed and experienced is real. Just as knowledge is tentative, so too are values.

Existentialism. This modern school of thought grew out of the work of the Danish philosopher Soren Kierkegaard, who believed that the central problem facing humanity is the ability to cope with its own existence. Individual freedom is viewed as being of primary importance. Since there are no absolutes, the individual is what he or she determines to become. One must choose what is essential and meaningful for oneself in this life and accept the consequences of one's choices.

Traditional and Contemporary Philosophies: Purposes and Roles

Traditional Philosophies

Ellis et al. (1991) also categorize two areas of educational philosophies which grew out of these aforementioned five: the traditional and the contemporary.

The traditional philosophies are based on perennialism and essentialism; the contemporary ones are based on progressivism, reconstructionism, and existentialism. As Ellis et al. have noted, there is no conscious attempt at this time to rank one philosophy as better than another. Each philosophy has its "intelligent, well informed, and thoughtful proponents" as well as its "ardent critics" who are just as intelligent, informed, and thoughtful. Included here are the general outlines of these two categories in terms of the purpose, the curriculum and method, the role of the teacher and the role of the school:

Perennialism. Perennialism has its philosophic base in the schools of Idealism and Realism. The basic purpose of the perennialist education is to help the student uncover and internalize the lasting truths. These universal and constant truths are the goal of education according to the perennialist. The training of both the intellect and the spirit are central. Robert M. Hutchins, a longtime proponent of the perennialist school, summarizes education's task in Ellis et al. (1991):

Education implies teaching. Teaching implies knowledge. Knowledge is truth. The truth is everywhere the same. Hence, education should be everywhere the same. (p. 104)

The curriculum for the perennialist is subject centered and draws heavily on the disciplines of literature, mathematics, languages, and the humanities, including history. The

method of study is the reading and discussion of the "great works" which in turn disciplines the mind. The teacher must be one who has mastered the discipline and is viewed as an authority whose knowledge and expertise are not to be questioned. The role of the school becomes one of training an intellectual elite who know the truth and will one day be charged with passing this truth to a new generation of learners.

Essentialism. Essentialism has its philosophic base in the Neo-Thomist school. This has been the most predominant educational philosophy throughout history and the modern essentialist movement actually developed in response to progressivism (which will be dealt with in the next section). Essentialism also draws upon the schools of Realism and Idealism. The major position was formulated by Professor William C. Bagley, regarded as the father of the essentialist philosophy. He believed that the major function of the school was to transmit the cultural and historical heritage to each new generation of learners. He is summarized in Ellis et al. (1991) in a discussion of his essentialist philosophy:

1. Gripping and enduring interests frequently grow out of initial learning efforts that are not intrinsically appealing or attractive.
2. The control, direction, and guidance of the immature by the mature is inherent in the prolonged period of infancy or necessary dependence peculiar to the human

species.

3. While the capacity for self-discipline should be the goal, imposed discipline is a necessary means to this end. Among individuals, as among nations, true freedom is always a conquest, never a gift.

4. Essentialism provides a strong theory of education, its competing school (progressivism) offers a weak theory. If there has been a question in the past as to the kind of educational theory that the few remaining democracies of the world need, there can be no question today. (p. 106)

The purpose of essentialist education is to pass on the cultural and historical heritage through a core of accumulated knowledge which has persisted over time and thus is worthy of being known by all. The curriculum is subject centered, with the elementary emphasis on basic skills. Mastery of basic facts and concepts of the essential disciplines is imperative. The role of the teacher is much like that of the perennialist. The classroom is very much under the teacher's influence and control. The role of the school becomes one of conserving and transmitting to the current generation the cultural and historical heritage thought necessary to make the student a contributing member of society.

Contemporary Philosophies

Progressivism. Progressivism has as its philosophical base the schools of Experimentalism and Pragmatism. As an educational theory, progressivism grew out of the pragmatist theories of people like John Dewey. Dewey viewed the school as a miniature democratic society with an emphasis on how to think rather than what to think. The basic underlying principles of progressivism as summarized by Kneller (cited in Ellis et al., 1991) are outlined below:

1. Education should be life itself, not a preparation for living.
2. Learning should be directly related to the interests of the child.
3. Learning through problem solving should take precedence over the inculcating of subject matter.
4. The teacher's role is not to direct, but to advise.
5. The school should encourage cooperation rather than competition.
6. Only democracy permits-indeed encourages- the free interplay of ideas and personalities that is a necessary condition of true growth. (p. 108)

The purpose of progressive education then is to give the individual the necessary skills and tools with which to interact with his or her environment - an environment which is constantly changing. The learning process should focus on cooperative behaviors and self-discipline. The curriculum is generally built around the personal and social

experiences of the students. Books are viewed as tools in the learning process rather than as sources of ultimate knowledge. The scientific methods of inquiry and problem solving are the generally accepted methods. The role of the teacher becomes that of a guide for the students in their problem-solving activities and projects. The role of the school is to become a living-learning laboratory, a working model of democracy - a microcosm of the large society.

Reconstructionism. The philosophic base for this school of thought is also Experimentalism and Pragmatism. Often referred to as social reconstructionism, this philosophy grew out of the progressive movement. George S. Counts was instrumental in forming the early thinking of this movement in 1932 with his work Dare the Schools Build a New Social Order?. Counts writes (in Ellis et al. 1991):

If the schools are to be really effective, they must become centers for the building and not merely for the contemplation of our civilization. This does not mean that we should endeavor to promote particular reforms through the educational system. We should, however, give to our children a vision of the possibilities which lie ahead and endeavor to enlist their loyalties and enthusiasms in the realization of the vision. Also our social institutions and practices, all of them, should be critically examined in the light of such a vision. (p, 110)

The reconstructionists believe that progressivists are

concerned only with the problems of society as it presently exists and that they do not go far enough in their efforts to improve society. They believe that what is needed in this age of rapid technological advancement is the reconstruction of society and the creation of an entirely new world order. The purpose of reconstructionism is to raise the consciousness of students regarding the social, economic, and political problems facing humankind on a global scale, and to instruct them in the necessary skills to solve these problems. The curriculum uses the organizing structures of the social science disciplines and the processes of scientific inquiry as the methods for working toward the solution of problems. The role of the teacher is to make students aware of the problems and then to make sure they have the necessary skills to work on the problems. The school becomes the primary agency for reform in society.

Existentialism. The philosophic school of Existentialism is the base for this viewpoint. The basic purpose of school, according to this position, is to enable each individual to develop his or her potential for self-fulfillment. A.S. Neill, a teacher in the rigidly disciplined Scottish schools, was led to found the student-governed Summerhill and was to become a major proponent of existential philosophy in education. Neill's belief in Freudian psychology was said to assist him in formulating his ideas about the ideal education. Ellis et al. (1991) note:

(Neill's) basic belief was in existential noninterference - freedom of choice and self-government; according to Neill, this was the best treatment for delinquency, for when faced with decisions, people choose what is best for them ... Neill believed that students would not abuse their freedom; rather, he felt that freedom was abused when it was cruelly withheld ... what Neill viewed as hypocritical moralizing was absent from the curriculum. Existentialism insists not that the teacher be "successful," but that the teacher be honest. Nevertheless, honesty leads to success, for if the teacher is honest with the pupil, trust is established ... thus the dialogue that is education rests on trust between persons, a trust that the teacher must earn by integrity and create with skill. (p. 114)

The curriculum in existentialism is not generally prescribed, but the students are encouraged, through a process of reflective thought, to pursue projects that will help them develop needed skills and acquire requisite knowledge. The teacher's role is to guide the learner and gently stimulate reflective thought through probing questions. The role of the school is to become a forum where teacher and student engage in dialogue to help clarify progress towards self-fulfillment.

An awareness is needed of these philosophies and their contributions to educational thought in order to understand

and contrast more fully the dominant paradigm - that of the factory model reflected in the systems approach or the instructional design approach.

A View of Humanity

Dobson, Dobson and Koetting (1985) in their book, Looking At, Talking About, and Living With Children: Reflections on the Process of Schooling have attempted to identify and contrast three philosophical and psychological profiles by separating them into three camps called Design A, Design B and Design C. They note that "the separation is quite possibly a direct reflection of whether persons are primarily concerned with doing to, for, or with young people" (p. 36). They further note that the three camps can be dispersed along a continuum which ranges from the the training of children to the education of children.

MacDonald (cited in Dobson et al., 1985, p.42) defines both, writing that "...training is the process of preparing a person to perform defined functions in predictable situations and education is the process of equipping an individual to perform undefined functions in unpredictable situations." Dobson et al. support this view adding that an educational program committed to training is based on the belief that humans are the sum total of their experiences and passive victims of their environments. On the other hand, education, at the opposite end of the continuum is committed to the idea that humans are active, goal-seeking

organisms eager to profit from encounters with the environment (p. 42).

They go on to explain that what people believe about humankind influences how they interact with others. Those who follow the Design A profile believe that humans are basically evil and that children need to be directed and controlled. "These people attempt to shape learners according to their values and teach children what they should know" (p. 42).

Design B people tend to hold a neutral belief about humans, but also tend to maneuver children toward predetermined goals. "Design B proponents begin with children where they currently are functioning and manipulate the environment so children have the best possible experiences based upon the adults' perception of what is best" (p. 42). These people encourage choice making, problem solving, creativity and autonomy.

Those who favor Design C think that humans are basically good, cooperative and interested in enhancing their uniqueness. In Design C, the children are accepted and given stability to interact with all others in the school setting.

The nature of society as an outgrowth of the nature of the humans who inhabit and create it is also influenced by these three designs. According to Design A, school is an institution which seeks to preserve the culture and maintain the existing social order. Design B views society as a

process in which the individuals participate and education should provide new ideas for planning for the future.

Design C proponents believe that improving individuals, not institutions, is the way to improve society. The school should concentrate on the development of freedom in the child.

This design model includes a number of other aspects besides the view of human nature just described. However, it is these views of humankind that help make distinctions among the various educational schools of thought, and while Dobson et al. (1985) caution that these designs are seldom found in pure form, they also hold that most schools are patterned after one of the three.

The Role of Schools

While it is important to understand the view of humankind that governs a school organization, one cannot discuss education without taking into consideration the different influences and the different demands placed on the schools and educators by the American public. In addition to being aware of the ideologies which have influenced the growth of the American educational system, it is also of utmost importance to have a general idea of what is perceived as the role of the school today. Assessment in the classroom is very much a product of the need to know what students are doing and how. This need to assess originates in the questions that ask what schools are doing

and how. Ultimately, according to Rowntree (1987):

if we wish to discover the truth about an educational system, we must look into its assessment procedures. What student qualities and achievements are actively valued and rewarded by the system? How are its purposes and intentions realized? To what extent are the hopes and ideals, aims and objectives professed by the system ever truly perceived, valued and striven for by those who make their way within it? The answers to such questions are to be found in what the system requires students to do in order to survive and prosper. The spirit and style of student assessment defines the de facto curriculum. (p.1)

This need to assess and to evaluate will be discussed in greater detail in Chapter Three. At this point the question is: what are schools for? Many people have attempted to answer this question in terms of the knowledge that is offered by way of the curriculum.

The Overt and Covert Curriculum

Michael Apple (1990) asks "in whose interests do schools often function today? What is the relation between the distribution of cultural capital and economic capital? What interests do the schools serve, those of the parents and children, or those of the teachers and headmaster?" (p. 59). Apple (1990) and others (Green and Sharp, 1975; Goodlad, 1984; Young, 1971; cited in Apple, 1990, pp. 173-

174) concur that the educational system has been dominated by a perspective that might best be called "technological", in that the major interest guiding its work has involved finding the one best set of means to reach pre-chosen educational ends (p. 44). "The overt and covert knowledge found within school settings, and the principles of selection, organization, and evaluation of this knowledge, are value-governed" according to Apple (1990, p.45). "The curriculum in schools responds to and represents ideological and cultural resources that come from somewhere. Whose meanings are collected and distributed through the overt and hidden curriculum in school?" (Apple, 1990, p. 46).

Goals, Functions and Aims

In an attempt to discover what the schools are for and who decides, John Goodlad (1979) notes that Americans are impatient today to talk about the fundamental issues pertaining to schooling. They want to know what kind of individuals the schools should seek to develop, what kinds of experiences young people should have in schools, and most of all what education is. In his discussion he posits three major questions concerning what education is: What are schools expected or asked to do? What do schools do? What should schools do? Goodlad goes on to make distinctions as to the use of the words goals, functions and aims. What the schools are asked to do he refers to as goals. What the schools are used for or do he refers to as function. What

the schools should do, the "ideal postulates", he refers to as aims. Concerning the function of schools, he comments:

Schools perform two sets of functions: 1) social functions not expressly stated as goals and 2) legitimated educational goals (as well as some educational functions not so legitimated). Both sets of functions take up time and resources, the non-educational ones sometimes consuming more than the educational. In appraising the role and performance of schools, however, we concentrate almost exclusively on their educational function - and usually on only a small part of it at that. (p. 8)

Goodlad also notes that education must be evaluated not just according to goal attainment but also according to the means employed. Or conversely, means must be judged by more than their contribution to predetermined ends.

Goals as a Socio-Political Process

In discussing the goals for schooling Goodlad says that "goals for schooling emerge through a socio-political process in which certain sets of interests prevail over others for a period of time" (Goodlad, 1979). These goals are what he calls:

client perceived wants and needs, professional determinants, pervasive interests of the citizenry in teaching a common culture, expectations and colleges and universities, and the economic interests of

business and industry. (p.44)

Goodlad goes on to categorize the goals that have emerged in the United States:

- 1) academic - early emphasis was on sufficient schooling to learn the principles of religion and the laws of the land (sometimes referred to as functional literacy);
- 2) vocational - readiness for productive work and economic responsibility;
- 3) social and civic - socialization for participation in a complex society; and
- 4) personal - the goal of personal fulfillment, which is a fairly recent development. (p. 44)

Out of these four general categories there emerged a set of 12 goals that Goodlad considers to be a "reasonably accurate and comprehensive summary of our verbal, and to a degree, our ideal commitment to goals for schooling" (Goodlad, 1979, p. 46). These goals are:

1. Mastery of basic skills or fundamental processes
2. Career education - vocational education
3. Intellectual development
4. Enculturation
5. Interpersonal relations
6. Autonomy
7. Citizenship
8. Creativity and aesthetic perception
9. Self-concept

10. Emotional and physical well-being

11. Moral and ethical character

12. Self-realization

One of Goodlad's conclusions concerning the goals, functions and aims for schools is that the question of what schools are for is usually settled in the socio-economic marketplace, not the schools (Goodlad, 1979, p. 57). Apple (1990) would agree when he observes that "we begin to see how a society reproduces itself, how it perpetuates its condition of existence through the selection and transmission of certain kinds of cultural capital" (p. 60). Additionally, he demonstrates that "education is a political process" (p. 60) and that the social and economic foundations upon which the educators act are not neutral.

The Language and Metaphors of Schooling

"Functions get established," notes Goodlad (1979, p. 57) by custom, by fiat, through legislative act, and by rule of the courts. They are perpetuated through mechanisms created for the conduct of schooling. One of those mechanisms is the language of metaphors.

The next section of this chapter will look at language and identify some of the metaphors used to describe the education of the children of America and the comments of some noted theorists concerning the philosophies out of which these metaphorical approaches have emerged.

Goodlad has said that there is a language of schooling

- a language noted by him and others that is largely scientific and mechanistic in nature. According to Schubert (1986, p.180) "the language that one uses has a great influence on both communication and on the way in which one views the world. Apple (1975), Friere (1970), and Wesker (1976) (cited in Dobson, Dobson and Koetting, 1985, p. 5) "suggest that language is not passive or neutral."

Metaphors As Shapers of Social Reality

"Educators," according to Dobson, et al. (1985) "invent words to use as tools and their perceptions become controlled by these creations." (p. 6). Metaphors are ways, through language, to understand what isn't already understood. The caution is not to confuse the metaphor with reality. "Language which is intended to explain or describe reality becomes reality. What can't be explained...is too often ignored and ultimately dismissed" (Dobson et al. 1985, p. 6). Dobson et al. (1985) have even suggested that the language used in the field can often be deterministic and can "encourage human encounters a priori. In other words, the language determines what is seen before we look" (p. 6).

Lakoff and Johnson (cited in Szatjn, 1992) suggest that human thought processes are metaphorical and that human conceptual system is metaphorically structured - that is, one concept is understood in terms of other concepts that are more natural or familiar. They further remark that "experiences take place within a background of cultural

pre-suppositions and that the fundamental values of a culture are coherent with the metaphors chosen for the fundamental concepts in that culture" (p. 36). They also predict that:

Metaphors may create realities for us, especially social realities. A metaphor may thus be a guide for future action. Such actions will, of course, fit the metaphor. This will, in turn, reinforce the power of the metaphor to make experience coherent. In this sense metaphors can be self-fulfilling prophecies. (p. 36)

Apple (1990) contends that "...models are for understanding, not necessarily for control." But in noting that the model often becomes, not the explanation, but the model by which the entire schooling system is governed, he comments that "we have yet to learn the dangers of appropriating models from disparate fields and applying them to education" (p.113). Apple also remarks that "the quite basic procedures of languaging and thinking that dominate education today give meaning (and latently prevent other forms of meaning from being seriously considered)" (1985, p. 122).

In discussing the necessity of examining the power of words and the language of metaphors, Dobson et al. (1985) wrote that:

Professionals must deal not only with what they see but with why they see what they see...The way educators

look at (perceive), talk about (language), and live with (experience) children is an area worthy of critical analysis" (p. 7).

The question then is: How do educators look at, talk about and live with children? What guides the perception and the language with which children are experienced in the classroom? Kliebard (1958) categorizes what he considers to be the prevailing metaphors into the following:

The Metaphor of Production. The curriculum is the means of production, and the student is the raw material which will be transformed into a finished and useful product under the control of the highly skilled technician.

The Metaphor of Growth. The curriculum is the greenhouse where the students will grow and develop to their fullest potential under the care of a wise and patient gardener.

The Metaphor of Travel. The curriculum is a route over which students will travel under the leadership of an experienced guide and companion. (p. 84, 85)

Dobson et al. (1985) explain that as a consequence of the narrowly limited vision of the constructs of intelligence and behavior used to describe children, "there have evolved essentially three sets of metaphors used in talking about children; military, industrial and disease" (p.8). They go on to cite Huebner (1963) who classified

values into five frameworks: technical, political, scientific, aesthetic, and ethical. However, they also note that "while none of these values is inherently destructive; the exaggerated dependence on some to the exclusion of others is dangerous" (p. 8).

Dominant Metaphor: The Factory
Metaphor

The dominant metaphor influencing curriculum and education in general as noted by Schubert (1986), Goodlad (1979), Apple (1990), Dobson et al. (1985) and others, is largely that of the factory model. Concerning this model Schubert (1986) observes that:

the dominant curriculum language (that of the theoretic, conceptual empiricist, or social behaviorist) reveals a world of persons as potential products who are forged on the assembly lines of schools and are judged by methods of quality control that utilize technical, quantitative jargon. (p. 180)

He also explains that "if the students do not conform to the factory model of growth, they are reshaped by the military model of control and obedience to authority."

Apple (1990) concurs in this assessment when he observes that students who do not fit the mold provided by the school are the focus of steps to correct the deviancy:

Punishment, rehabilitation, therapy, coercion, and other common mechanisms of social control are things

that are done to him (the student), implying that the causes of deviance reside within the person to whom the label has been attached, and that the solutions to the problems that he presents can be achieved by doing something to him. (p. 135)

This becomes even more important later on in this study when the issue of labeling is discussed as an outgrowth of assessment practices.

Szatjn (1992) is critical of current attempts which suggest that American schools should follow the Japanese industries as models and should be managed as enlightened corporations. Szatjn also, in summarizing previous articles concerning the Total Quality Management philosophy of W. Edwards Deming, noted that they "covertly implied or overtly stated" the necessity of a paradigm shift in education. This shift would apply Deming's principles to education, a situation that Szatjn considers, merely a change in metaphor, not a paradigm shift. Moreover:

the metaphor proposed is not much different from the one people are now criticizing schools for. Changing the school is a factory metaphor for the school is an enlightened corporation one is just updating the business metaphor. We are still using economic principles and vocabulary to express educational ideas. We are still allowing economy and production to shape and determine our understanding of education. We are still seeing students as raw materials to be processed

in the most efficient way. (p. 36)

In spite of a number of reforms in education, it is still the factory model or some form of it which influences thinking and continues to influence the way school is conducted today.

The Conflict of Articulation

Schubert (1986) points out "anytime that we "alter our mode of educating others, we indirectly influence the character of some of our other assumptions about education". Implementing changes possibly will conflict with assumptions that guide other ideas which are popular. "If the accepted model is behavioristic and most of our curriculum is humanistic, there is conflict." Along with Schubert's caution that the "assumptions that we report possessing may not be the ones that actually guide us" it is worth noting that Goodlad (1984) has also suggested that what the schools say they do and what they actually do may be two different things. Even to reflect on a philosophy of education as suggested by Apple (1990) does not mean that it is practiced. Schubert (1986) points out that "to articulate philosophy does not necessarily mean that we practice it successfully." Apple (1990) concurs when he says:

One of the difficulties in seeking to develop new perspectives is the obvious and oft pointed to distinction between theory and practice or, to put it in commonsense language, between 'merely' understanding

the world and changing it. (p. 103)

This conflict between emerging modes of thought and the attendant practice is referred to as crisis by Thomas Kuhn (1970) and is a factor in causing paradigm shifts - shifts in one's world view and the emergence of competing paradigms to take the place of the existing one. Goodlad (1979) calls attention to two significant signs of being at the end of an era even if the outlines of a new one are far from clear:

First, assumptions about our schools previously unquestioned or questioned only by radicals have begun to come in for more serious popular questioning ...

Second, the less tenable long-established assumptions appear to be, the more intense the ceremonial rain dances performed by those who fear the personal consequences of new approaches. That is, threatened groups and individuals try harder to do what gave satisfaction before, however inappropriate and outworn such behaviors may be. (p. 67,68)

Since the existing paradigm is one of behaviorism and control, and current studies are proving that the tests, and quantification that go along with this paradigm are not meeting the goals and aims of the educational system, then it becomes necessary to examine alternatives. As mentioned at the beginning of this section Goodlad (1979) has commented that:

the ultimate test of what schools are for is what they do. What they actually do may bear little relationship

to what they are officially asked to do ... An anomaly begins to emerge ... (p.7)

This anomaly that Goodlad refers to is another factor leading to crisis as outlined by Kuhn (1970). This crisis eventually is resolved in a paradigm shift. Dobson et al. (1985) suggest the need for "alternative constructs and language for viewing and talking about children in order to enhance their educational living experiences." They, along with a number of others (Eisner, 1979; Leonard, 1972; Macdonald, 1968; cited in Dobson et al., 1985, p. 9) suggest that new constructs and a new language can be found in, for instance, the aesthetic value system. This value system provides a challenge to the scientific, systems management, factory approach that has for so long dominated educational practice.

In conclusion, this chapter has taken a brief look at the philosophical foundations which have influenced educational thought through the ages. In addition, the chapter looked at the existing metaphors and language which have controlled pedagogy and which pervade educational thought and practice. As was indicated in the introduction to this analysis, one purpose of this study is to provide an examination of an alternative to the existing paradigm in the hope that it might become a useful alternative to the current normative testing and grading that now exists. Chapter Three will take a critical look at the current assessment practices now in use in the American school.

CHAPTER III
CONCERNS, CRITICISMS, AND DEFENSES
OF ASSESSMENT

Assessment as Control

This chapter will address concerns of, criticisms of, defenses of and an alternative form of assessment. Chapter Two attempted to present an overview of some of the major philosophies which have influenced American educational practices over the past 200 years. It was shown in Chapter Two that a number of metaphors have emerged as descriptors of the prevailing methods of providing schooling. These metaphors and the language that is used in association with them represent differing orientations to curriculum theory, research and practice. By far the most dominant of these metaphors has been that of industry. This orientation involves a heavy reliance on quantifiable data, on scientific approaches of observation and objectivity and on control. Kliebard (1972, cited in Koetting, 1990) observes that in this metaphor of production:

the curriculum is the means of production, and the student is the raw material which will be transformed into a finished and useful product under the control of a highly skilled technician. The outcome of the

production process is carefully plotted in advance according to rigorous design specifications, and when certain means of production prove to be wasteful, they are discarded in favor of more efficient ones. Great care is taken so that raw materials of a particular quality or composition are channeled into the proper production systems so that no potentially useful characteristic of the raw material is wasted. (p. 84)

One of the most influential books on curriculum and practice was Ralph Tyler's Basic Principles of Curriculum and Instruction (1949). This paradigm, or "conceptual lense through which curriculum problems are perceived" (Schubert, 1986, p. 2) was first written as a syllabus for a course taught by Tyler at the University of Chicago. It found its way into book form and has become one of the most widely cited curriculum books, having been translated into at least ten languages (Schubert, 1986, p. 171). In this book Tyler (cited in Schubert, 1986, p.172) identified four questions that should provide the parameters for curriculum study and uses the questions as his chapter titles as follows:

1. What Educational Purposes Should the School Seek to Attain?
2. How Can Learning Experiences Be Selected Which Are Likely to Be Useful in Attaining These Objectives?
3. How Can Learning Experiences Be Organized for Effective Instruction?
4. How Can the Effectiveness of Learning Experiences Be

Evaluated?

Schubert (1986, p. 171) calls attention to the educational reputation that Tyler had already developed as the director of evaluation for the Eight Year Study among other things. He indicates that Tyler in a 1980 interview:

openly acknowledged that his 1949 rationale was an attempt to summarize and synthesize what had been said earlier by other curriculum writers such as Franklin Bobbitt, W.W. Charters, John Dewey, Boyd Bode, Harold Rugg, and Henry Harap ... yet the manner in which Tyler handled the questions, by pointing out criteria and principles for decision, makes his work unique. (p. 172)

Schubert (1986) also notes that in the years that followed, many curriculum writers, enticed by technology, translated the intent of Tyler's Rationale into a how-to manual.

Existing Assessment Practices

The mode of curriculum inquiry that grew out of this fascination with the Tyler Rationale was influenced in large part by the behavioral sciences which were manifested in the disciplinary approach of subject matter areas. It was only fitting that the way to evaluate the curriculum packages that developed as a result of this rationale was through assessment that was empirical, analytical, behavioral and objective. As was reported by Broudy: "Quantitatively expressed data were seen as the prime source of truth and

the only avenue to credible reports of accomplishments" (cited in Schubert, 1986, p. 173). The empirical-analytical mode of research became the governing mode for the conduct of educational research and every area of curriculum, including evaluation or assessment.

Schubert (1986) describes the emergence of curriculum evaluation as having moved through several stages that began with an emphasis on grading, marking and judging and has proceeded to become a specialized system of measurement. The Eight Year Study (1933-1941) is said to have expanded the notion of evaluation beyond mere measurement to become the focus of accountability measures in today's schools. Numerous issues are the subject of much debate as outlined by Schubert (1986):

- (1) the relative value of quantitative and qualitative methods and the possibility of integrating the two;
- (2) the worth of theoretic and practical epistemologies, which invokes the issue of outside expertise versus locally developed evaluation;
- (3) the relative value of different disciplinary metaphors that should undergird evaluation, which principally refers to scientific technology or artistic criticism but may pertain to legal, journalistic, anthropological, and literary images of evaluation; and
- (4) conflicts over personal and political aspects of evaluation and the kinds of critical discourse that can help to unravel them. (p. 278)

Assessment as a Definition of Curriculum

With the above in mind, this chapter will look at the characteristics currently deemed desirable of evaluation procedures, at the existing assessment practices and their reliance on the standardized test, at the advocacy and criticism of these practices, and at an aesthetic alternative. As Rowntree (cited in Satterly, 1989, p. 37) has commented, "the problem of assessment is not only - nor even primarily - one of the development of better techniques. Far more fundamental questions are involved." Chapter Two quoted Rowntree (1987, p. 1) as observing that the truth of an educational system can come from looking at its assessment procedures and that the "spirit and style of student assessment defines the de facto curriculum."

That spirit and style which pervades the American school system has changed in an effort to make testing more efficient, manageable, standardized, objective, easier to administer, and less costly according to Madaus and Tan (1993, p.55). "In the 18th century, the oral disputation (which was) aimed at assessing universal rhetorical skills was supplemented by the written exam to more efficiently assess mathematics ... " (Madaus & Tan, 1993, p. 56). They also note that in the interest of making exams more uniform, more efficiently administered and more easily compared, Horace Mann introduced the written essay exam in 1845. This was followed in the first part of the 20th century with the short answer exam. Samuelson (cited in Madaus & Tan, 1993,

p. 56) is said to have named Frederick Kelly as the inventor of the multiple choice test item. The studies of Starch and Elliot (cited in Madaus & Tan, 1993, p. 56) are said to have led to the development of short, easily scored test items. Part of the reason was the result of findings which showed that the marks assigned to essay questions were very unreliable. Another part of the reason was in response to the growing requirement of the scientific management movement's requirement that children's achievement be tested to measure a district's efficiency (Callahan, cited in Madaus & Tan, 1993).

The onset of World War I required that a number of recruits be tested quickly and cheaply with the result that the Otis developed a group-administered IQ test. Later, in 1926 the College Entrance Examination Board opted for a multiple-choice format partly due to the cost of scoring and partly to allow for a greater variety in the test (Angoff and Dyer, cited in Madaus & Tan, 1993).

In 1930, an estimated five million standardized educational tests were administered annually according to Strenio (cited in Madaus & Tan, 1993, p. 60). By 1990, it was estimated that each year elementary and secondary students were taking 127 million separate tests as part of standardized test batteries. The continued growth in testing is reflected in the growth of state mandated testing programs, the growth of test sales, and the growth of the number of column inches devoted to citations on testing in

the Education Index.

Madaus and Tan (1993) consider the conclusion to be obvious: educational testing has expanded dramatically over the last 50 years in terms of both volume and societal importance. They also note that the uses of test results have also changed. According to them the National Commission of Testing and Public Policy noted in 1990 that dramatic growth in testing since the 1950s was coupled with the trend of greater reliance on test results to make critical decisions about children, such as:

- . Entry to and exit from kindergarten
- . Promotion from grade to grade
- . Placement in remedial programs
- . Graduation from high school

Further, "there was a dramatic increase in the use of students' scores to hold school systems, administrators, and teachers accountable" (p. 65).

Michael Apple (1990) echoes Rowntree when he reiterates:

Social and economic values ... are already embedded in the design of the institutions we work in, in the "formal corpus of school knowledge" we preserve in our curricula, in our modes of teaching, and in our principles, standards and forms of evaluation. Since these values now work through us, often unconsciously, the issue is not how to stand above the choice. Rather, it is in what values I must ultimately choose. (p. 9)

Schubert (1986) also recommends that the student of evaluation should be a student of values since "evaluation is an axiological problem, not merely a technical one" (p. 285).

Desirable Characteristics of Assessment

Even though a discussion of educational evaluation should not merely be a technical one, it is necessary to call attention to some of the technical aspects of evaluation and assessment. This section begins with a non-technical discussion of the characteristics desirable of all currently accepted evaluation procedures. Art Burke (1991) lists three areas of concern: objectivity, reliability and validity. He maintains that an evaluation procedure is objective to the degree that different scorers of the same material obtain the same results. An example of a highly objective evaluation procedure, according to Burke (1991) is a multiple-choice standardized test with an answer sheet scored by a computer. An example of a highly subjective evaluation procedure is the essay, where different raters may use different scoring rules.

Reliability. Reliability refers to the consistency of results produced by an evaluation method. All evaluation results are affected by measurement error; that is, error due to the presence of factors which affect test performance but are extraneous to the trait being measured. For example, Burke (1991) acknowledges that a student's score on

a mathematics test could be affected because:

the test contained material studied the night before, because the student was ill or fatigued at the time of testing, because the student hit a jackpot of lucky guesses, or because of a host of other circumstances not related to his or her real level of math achievement. (p.2)

Reliable evaluation results are those in which the influence of measurement error is minimized.

Validity. Validity refers to the adequacy and appropriateness of interpretations about the results of an evaluation method. Burke (1991, p. 2) notes that "for a test to be a valid measure of achievement in history, scores have to depend on knowledge of history, not on reading ability or test-taking skill." It is important to understand that validity is situational: An evaluation procedure may be valid for some purposes and invalid for others. As Burke (1991) points out:

For example, a test valid as a measure of achievement in third-grade history might not be valid as a predictor of achievement in fifth-grade history. Lastly, strictly speaking, it is interpretations or uses of evaluation results which are validated, not the evaluation procedure itself. (p. 3)

Criterion-Referenced versus Norm-Referenced.

Assessment can also be characterized based on whether it

falls under criterion-referenced or norm-referenced evaluation. Criterion-referenced evaluation represents categorical judgments in which a fixed set of standards are adopted and the achievement of each student is judged against these standards. Criterion-referenced testing is the establishment of specific measurable performance standards in the forms of goals, objectives, outcomes, and/or performance standards which are developed by the instructor and given to the student in advance of any instruction. Scores and/or grades are given to the individual student and are not dependent on comparisons to other students. A number of techniques are used to insure that the student has the greatest amount of success possible.

In contrast, norm-referenced evaluation represents comparative judgments in which the performance of one student is compared with the performance of other students by ranking students in order of performance. The student is graded in terms of how he or she ranks in some norm group, usually the student's classmates. Norm-referenced evaluation is often criticized because it represents potentially destructive competitive aspects, it represents learning relative to someone else's learning, it discourages those students who are below the mean, and it prevents students from progressing if their entire group shows progress.

Summative versus Normative. Another aspect of evaluation that was introduced by Scriven (cited in Schubert, 1986, p. 265) was that of formative and summative evaluation. Summative evaluation refers to the final appraisal of a program. Formative evaluation provides information about the program while it is still in progress. The information provided acts as a guide for the program.

The Values Implied in Assessment

Rowntree (1987) writes that assessment involves putting a value on something (usually in financial terms). He also acknowledges that, while this type of definition is not centered in educational assessment, it does fit with what many teachers think of as "essential components of assessment, viz the assigning of numerical marks or letter grades, and the ranking of students in order of preference or relative achievement. More basically, Rowntree notes:

assessment in education can be thought of as occurring when one person, in some kind of interaction, direct or indirect, with another, is conscious of obtaining and interpreting information about the knowledge and understanding, or abilities and attitudes of that other person. To some extent or another it is an attempt to know that person. In this light, assessment can be seen as human encounter. (p. 4)

The Purposes of Assessment

Rowntree (1987) also lists six main reasons commonly advanced for formal assessment of students. First he notes that one very common purpose of assessment is for the selection of candidates for various opportunities or careers. Students are selected for advancement or for non-advancement in many cases. Dave and Hill (cited in Rowntree, 1987, p. 20) talk about the examination system as forming the basis of a type of caste system." Apple (1990) affirms this when he writes:

schools engage in anonymizing and sorting out abstract individuals into preordained social, economic, and educational slots. The labeling process, thus, tends to function as a form of social control. (p.126)

A second purpose of assessment closely related to the first, is that of maintaining standards. The clientele of this information, employers and colleges, look for assurance of "quality control". The student is of secondary importance. Rowntree notes that it is hard enough to get the education community, itself, to agree on the standards, much less to agree on just what standard the student may have attained. It has been noticed by several (Miller, 1967; and Mills, 1972; cited in Rowntree, 1987) that even though the quality of students entering the University of California at Berkeley increased considerably, as determined by three different pre-entry criteria, their grade point averages remained the same, suggesting that the "standard of

'output' is maintained despite an apparent improvement in the standard of 'input'" (p. 22).

A third purpose of assessment is said to be for the motivation of students. Rowntree (1987) notes that for every student who "confesses himself to be in need of a constant prod from assessment there will be another who claims to be distracted and enervated by it" (p. 22). It must also follow, that assessment used as a tool by the teacher to structure and legitimize the curriculum, could be defined as an instrument of coercion. The line between coercion and encouragement depends largely on the intentions and perceptions of the teacher and student and the relationship between them.

A fourth purpose of assessment is to provide feedback to the students. A student receives this feedback in the form of grades, marks, rank, and so on. Rowntree (1987) admits that this sort of feedback provides little unless the student is aware of the performance of the reference group (in the case of norm-referenced grading). The grade or mark is the least useful form of assessment according to Sassenrath and Garverick (cited in Rowntree, 1987). The grades or marks are non-specific and tell the student neither the merit of his work nor do they provide any feedback for improvement.

Rowntree also calls attention to the fact that feedback can come from a number of sources including texts, the materials with which he or she works, the other students,

and, of course the teacher. The fact that the teacher can provide feedback in a number of ways from grades to verbal comments to gestures and facial expressions is supported by the 1970 studies of Snyder (cited in Schubert, 1986, p. 105) who recognizes the "hidden curriculum" as being implicit demands that often run counter to explicit aims. For example, although a school may encourage a student to take harder "college preparatory" courses, that school may also provide such rewards for good grades that a student feels penalized by taking a more rigorous course and making a lower grade than a fellow student in an easier course. Rowntree (1987) contends that the student should become increasingly capable of providing his or her own feedback and become less dependent on the standards, strategies and validation of others.

A fifth purpose for assessment is to provide feedback for the teacher. This enables a teacher to identify where he or she has failed to adequately explain a new concept, confused an issue, given insufficient practice, and so on. It is often argued, however, that the current use of standardized assessments and other externally marked examinations provide feedback too late to be of much benefit to the teacher in modifying strengths and weaknesses.

Preparation for life, is the sixth purpose of assessment. Rowntree (1987) challenges the comment that assessment prepares students for "real life" and argues that there should be no distinction between one's educational

career and real life. He agrees that much of the informal assessment (in the form of approval and criticism) that goes on in school and college is related to the informal assessment that goes on in the rest of life. He questions, however, the statement in a 1971 article by Brian Cox (cited in Rowntree, 1987, p. 29) which says: "All life depends on passing exams ... To create an education system without examinations is to fail to prepare children for the realities of adult life." Rowntree (1987) maintains that life outside of education is not really like that and that "with the exception of the civil service and armed forces, most people seldom ever again meet the experience of being tested or examined on a prescribed syllabus for the purpose of being graded and ranked and chosen" (p. 29). He also maintains that assessment in industry and the professions is generally "informal, diffuse, ad hoc and continuous. It is based largely on the person's track record over a period of time and in fulfilling his duties rather than on what he can write about something at a given point in time" (p. 29). The question is, in a system that copies the factory and industry as a model, why is it that this more informal and more authentic manner of assessment is not also copied?

The Side-Effects of Assessment

Rowntree (1987) goes on to examine what he considers to be the side-effects of assessment. In conceding that the most well-intentioned act often produce results other than

were intended, he looks at eight aspects of assessment that he feels can be held accountable for certain side effects:

1. The prejudicial aspects of assessment.
2. Students' knowledge of assessment.
3. The extrinsic rewards of assessment.
4. The competitive aspects of assessment.
5. The bureaucratic aspects of assessment.
6. The nature of specific assessment techniques.
7. The giving of grades.
8. The reporting of assessment results.

The Prejudicial Aspects of Assessment. "Students are affected by assessment even before they themselves are assessed" according to Rowntree (1987, p. 36). Whether through previous experience or through knowledge of the "normal" stages of development, teachers plan a course of study based on this experience and knowledge. The problem arises when the student does not fit the norm. As Rowntree (1987) points out, "the danger of harmful side-effects arises only when such general assessments derived from other children are pursued in the face of contradictory evidence from the particular children we are working with."

This calls attention to one ever present side-effect of assessment - the prejudicial use of stereotypes. This is referred to as "labeling" in Dobson, Dobson and Koetting (1985) who wrote "labeling is a process whereby one human agent or group makes a value judgment about the appropriateness or inappropriateness of another's actions,

thoughts, or being" (p.20). Michael Apple (1975) charges that "clinical terms used to label children in schools tend to function as a form of social control; that is, the sorting of children into preordained social, economic, and educational slots" (cited in Dobson, et al., 1985, p. 20). Dobson et al. (1985) contend that labeling is a great industry and that once a person is trained in labeling, then it is necessary to "find or create individuals to fit those labels in order to maintain employment" (p. 20).

Another aspect of this problem is the tendency to over-generalize; that is, an early evaluation of the student's personality or work, whether favorable or unfavorable, will become the evaluation to other aspects. Brophy and Good (cited in Rowntree, 1987) say that some teachers are unable to respond to success when they are expecting failure and of course the alternate holds true. Innacurate and inflexible assessments may act as "self-fulfilling prophecy" as will be noted in the next section.

Students' Knowledge of Assessment. The idea of this side-effect may be found in Heisenberg's "Uncertainty Principle" which refers to the uncertainty attached to observations or measurements of something when the very act of observing or measuring may alter the thing being observed and measured. Once the student knows his or her behavior is being observed and measured or assessed, then he or she may change that behavior in some way (Rowntree, 1987). An even more potent side-effect of this knowledge is that of self-

fulfilling prophecy, an idea first discussed in 1948 by Robert Merton as "a false belief, the expression of which starts off a chain of events that makes the initial belief come true" (cited in Rowntree, 1987, p. 42). In other words, believing something is true will often make it true, especially if, as Rowntree notes, many people act on that belief (1987, p. 42).

Although follow up studies have not replicated Rosenthal and Jacobson's famous 1968 Pygmalion experiment (cited in Rowntree, 1987), there seems to be little doubt that teacher-expectations can affect pupils' attitudes and achievements, especially when the students are aware of the expectations. In this experiment, teachers were told that certain children (actually randomly selected) were late bloomers who could be expected to make significant gains during the year. At the end of the year these children had, indeed, made bigger gains than their classmates and it seemed that the imaginary difference had become real and the prophecy fulfilled. When students are aware of how they are being assessed they may be affected in their self-esteem, receptiveness, and level of aspiration according to Rowntree (1987, p. 44).

The Extrinsic Rewards of Assessment. Learning should be its own reward, but for some it is and for others it is not. Sociologist Talcott Parsons (cited in Rowntree, 1987) asks whether the student regards his learning as "expressive (valuing it as an opportunity to express and enlarge his

capabilities) or as instrumental (valuing it as a means toward the satisfaction of goals external to itself)" (p. 44). Rowntree argues that too many students are encouraged to regard learning and education as instrumental rather than expressive. Too often, the students want "the certificate" or "the degree" more than or instead of the learning that supposedly is signified by that certificate or degree. Desiring the approval of others is not necessarily detrimental to learning, but it can be when gaining the approval of others is the main reason for learning.

The Competitive Aspects of Assessment. As if the effects of learning for the sake of extrinsic rewards are not bad enough, what happened when there are not enough of these rewards to go around? The side-effects are then worsened by competition. As Robert Wolf declares (cited in Rowntree, 1987, p. 51) "The Pythagorean theorem does not flicker and grow dim as more and more minds embrace it." Learning and knowledge is a "free commodity". It is when one thinks of knowledge as approved by way of GEDs, SATs, grades, degrees, and so on that the supply is no longer unlimited. For one person to get more, another must get less. "Many assessment systems are competitive in that the extrinsic rewards they offer are in short supply and each student who wants them is asked to demonstrate that he is more deserving than others, or others less deserving than he is" (Rowntree, 1987, p. 51).

The Bureaucratic Aspects of Assessment. Assessment is considered bureaucratic when it is impersonal. First, one is not able to identify the assessor and second, the assessor does not regard the person being assessed as an individual. The assessment is done for efficiency and to satisfy a preconceived standard. In addition, assessment is big business. Rowntree notes that the Education Testing Service, whose battery of tests and examinations helps control entry to the U.S. meritocracy, as long ago as 1974 had an annual income of more than 50 million dollars according to Rein (cited in Rowntree, 1987, p. 59).

The bureaucracy, associated with assessment, stems also from standardization. As Rowntree protests "to treat people equally is not necessarily to treat them fairly" (p. 60). Yet another observation cited by Rowntree (1987, p. 61) is that of Thorstein Veblen who identified that even in 1918 industrial leaders were the real clients of the schools and that the pressure for exams and grades arose out of their need for a bureaucratically efficient estimate of graduates' usability. A number of authors (Apple, 1990; Glasser, 1986; Eisner, 1985b; Grant, 1991; and Sund and Trowbridge, 1974) have supported in their writing Rowntree's claim that "curriculum follows the examination" (1987, p. 61). That is, few schools seek to develop curricula that address the individual needs of their students but choose instead to meet the standard of whatever test is required by their school.

The Nature of Specific Assessment Techniques. There are many ways of coming to know a student such as conversation, observation, multiple-choice tests, essays, examinations and so forth. As Grant Wiggins, an advocate of "tests worth taking", warns: when it comes to any kind of testing, one size doesn't fit all" (cited in Brandt, 1992, p. 35). Side-effects occur when the reliance is too heavy on any one method of assessment. Any technique may cause side-effects if it is over-applied or it it is inappropriately applied.

The Giving of Grades. Much of the current criticism in the assessment field is aimed at the "grading system". Rowntree, however, fails to find a group of side effects particularly attributable to grades themselves (1987, p. 68). Most of the side-effects are already associated with other aspects of assessment. Rowntree does suggest that grades are to be blamed more for what they don't do than for what they do. What they don't do is tell all that is known about a student. Dobson et al. (1985) wonder how teachers can assess learning for every student by using only five letters of the alphabet.

Dreikurs, Grunwald and Pepper (1971) are cited by Dobson et al. (1985) as contending that good grades are motivating, but only to the students who are already motivated and already getting good grades. The less motivated students pay no attention to them and Dreikurs et al. conclude that grades "... are neither needed nor

effective" (p. 23). Hargis (1990) echoes this when he explains that success is fundamental to achievement. Lack of success, according to him, means a lack of achievement. Failing grades are indicative of the failure to provide success for most students who receive them. Grading will not motivate the low-achieving student who is not capable of more. That child, instead, will be demoralized. Evans and Glasser (cited in Hargis, 1990, p. 3) point out that "all you learn from failing is how to fail".

The Reporting of Assessment Results. Nearly every time a newspaper is opened one can find some type of assessment results. The papers will carry long lists and charts listing schools, grade levels and gains or losses in SAT or some other achievement test score. Student assessment is reported to a number of people including the student, other teachers, other students, the parents, potential employers, and so on. While no one disputes that the student and perhaps the parent have a right or need to know assessment results, the question becomes one of the right to privacy when it begins to involve others. Teachers may use reported information to mis-select or mis-teach a student and employers may use the information to mis-employ a student when the results of assessment are not reported in a form that yields reliable and relevant information.

A Defensive Rejoinder Concerning
Assessment

David Satterly (1989) provides in his book, Assessment in Schools, an apparent rejoinder to the arguments posed by Rowntree's 1987 study. He acknowledges that the theories, assumptions and practices associated with assessment deserve to be continually critically scrutinized. Neither the objections nor the replies completely address every aspect of the arguments but simply serve as a way to increase awareness of the issues concerning assessment and its effect on the individual, on the institution and on society.

For the sake of clarity, the argument will be given followed by the rejoinder that Satterly provides to it:

Assessment: A Political Activity. Assessment is a political activity which preserves the social order in society. Assessment is the way schools perpetuate the existing hierarchical structure of society and results in the application of labels which determine a child's future opportunities.

Satterly replies that it is over-simplification to argue that assessment is a fundamental cause of of the value system held by society. He also notes that this objection refers mainly to norm-referenced testing and not so much to criterion-referenced testing. In addition, he notes that there are some areas requiring special skills that just do require assessment. He points out also that assessment is

required as a means of justifying the existence of the educational program to those who fund it. He argues:

That the result of an assessment could be used to help perpetuate a model of society which a teacher holds to be undesirable is not an argument against the positive use of that same result to inform both teachers' and pupils' efforts to learn. (p.18)

Assessment and Class. Assessment favors children of the middle class, is to the disadvantage of other groups, and is too often interpreted as the result of differences in innate potential.

Satterly replies that this argument applies most to intelligence testing and replies that in this area there is concern about children from minorities who tend to do poorly. He counters with the fact that it is only through the use of these tests that children who are unable to profit from the educational program offered by the school are identified and hopefully given alternatives.

Assessment and the Cognitive Area. The comparative ease of assessment of the cognitive objectives (those associated with remembering, with reproduction of material and with the solution of problems) leads to a lack of attention to the more important objectives which are much more difficult to evaluate.

Satterly replies that there is some justification for this argument, but he argues that the more abstract the

objective the more difficult it will be to access. The objective is not only difficult to define in measurable terms, but it may also be simulated by the learner to a greater extent than would be possible with a cognitive objective.

Assessment and Expectancy. The results of assessment have an uncanny knack of being self-fulfilling. Assessment can not only reinforce a pupil's picture of self, but can be a part of the formation of that picture early on. Teachers also build up expectancies about pupils with the result that students perform the way they are expected to perform.

Satterly concedes that this often happens, but discounts the harmful effects by suggesting that teachers be aware of this phenomenon and counter it by using more criterion-referenced testing and become more aware of the influence of assessment data on their attitudes.

Assessment and School Curricula. Published forms of assessment - such as standardized tests - mold school curricula and inhibit new developments. The adoption of standardized tests tend to define the objectives of teaching and the teacher is, as Apple (1990) discusses, "deskilled" by the external assessment.

Satterly concurs that the determination of a school curriculum based on the content of a test reflects the misuse of the test. He contends that the school should choose the test based on the ability of the test to assess

what is already a part of the curriculum. In addition, teachers should develop tests which match their own objectives and which call for more attention to types of assessment which demand more than recall from the students.

Assessment and the Knack of Taking Tests. Assessment encourages the pupil to develop the styles of thought or intellectual 'tricks' required by tests and, therefore, inhibits the development of other skills.

Satterly argues that these objections conceive of assessment narrowly and overlook the range of instruments now available for assessment of all types. He also suggests that all children are given practice in test-taking when norm-referenced test are to be used. Again, he challenges the teacher to be the force in minimizing this objection.

Assessment and Role-Relationship. Assessment inevitably takes place in a role relationship. This is antithetical to a truly educational setting where encounters between teachers and pupils are interpersonal. Any assessment is made by someone about someone. The act of assessment often prevents the type of relationship that should be a part of the teacher-student encounter.

Satterly responds that this criticism is a fundamental criticism of schooling and not confined to assessment only. It is his belief that the personal encounters can take place in discussing the results of the assessment and he argues that there is nothing inherent in assessment which should

prohibit the establishment of the teacher-pupil relationship advocated by humanistic psychologists.

Assessment: An Invasion of Privacy. Many types of assessment are an invasion of the privacy of an individual. Educational assessment results are often misused. Only those types of assessment that deal with learning or later achievement should be permitted. Measures of personality, opinion, personal values, attitudes and background should be excluded.

Satterly admits that the responsibility is on the test user or the teacher making the assessment to make sure that the information gathered is useful to the child. He encourages the use of judgment and care to ensure that the schools avoid the charge of invasion of privacy.

Assessment Is Unreliable. All assessment - especially using tests - is unreliable and predicts imperfectly. If it is necessary to claim an often large margin of error by way of standard deviations and so forth, then the user is forced to question the cost-effectiveness of the program. Additionally, if these error-ridden scores only marginally improve predictions over chance predictions then why demand any prediction at all?

Satterly points out that decades of work in measurement theory and test construction have sought to reduce the unreliability of tests. He challenges that it is preferable to use test scores to quote the best estimate that can be

given of the size of the error than it is to rely on the subjective judgment of teachers who also make errors. He agrees that there is not justification for the use of tests which are consistently found to be invalid.

Satterly's conclusions are that an enormous amount of time and effort is expended in activities associated with assessing children. He acknowledges that there is often a great deal of stress associated with assessment. While he considers the questions about assessment to be worthy of discussion, he also remarks that "ideologies and the social contexts of the time influence the degree of acceptability of any theory or set of practices" (Satterly, 1989, p. 36). He acknowledges the "concern of many for what they see as disturbing trends in a modern industrialized society ... " (p. 36). In addition, he acknowledges:

... an increasing distrust of "science", there are suspicions of the widespread use of computerized records and fears that human being may be "reduced to numbers" by tests and records. To many people these trends imply a reduction in human freedom. Others fear that too much attention is paid by the education system to only a small portion of the abilities of children - usually their quantitative and verbal skills - with the danger that the complexity of human potential is overlooked. (p. 37)

Trivialization and Deskillling

Another disturbing trend briefly touched on by both Satterly (1989) and Rowntree (1987) is the tendency of current forms to assessment and grading practices to encourage trivialization and deskillling of the teacher.

This problem stems from the reliance on tests and curriculum packages and the reliance on objective scoring and assessment. Ease of testing becomes the criterion for curriculum content according to Sears and Marshall (1990). Madaus asserts that the curriculum is reduced to preparing for the exam and that schools should resist the tendency to standardize the curriculum through the use of easy-to-administer tests (cited in Sears & Marshall, 1990, p. 202).

McNeil (cited in Sears & Marshall, 1990, p. 203) designates four strategies used to trivialize the curriculum. These strategies include fragmenting knowledge into lists and facts, mystification of knowledge implied to be too difficult to master, omission of controversial, anomalous material, and defensive simplification to avoid student resistance to what they perceive to be difficult.

This trivialization goes on according to Sears & Marshall (1990) as a result of mandated curriculum testing practices, bureaucratic operation of schools and the lack of power for teachers to influence curriculum policies. Another result of this trivialization is the deskillling of teachers. Shore and Friere (cited in Sears & Marshall, 1990, pp. 201-202) summarize the effect of a "teacher proof"

curriculum that does not allow teachers to design content and tests based on student differences. Learning is designed, monitored and measured by managerial models of teaching rather than professional ones. Apple (cited in Sears & Marshall, 1990) has:

identified the loss of control over decision making by teachers and other local educators when the conception and execution of curricular and instructional plans are separated from each other. The former is done higher in the hierarchy, and the latter is left to the teachers - thereby deskilling teachers, that is, reducing their tasks to ones that can be carried out with a minimum of professional knowledge and judgment. (p. 201)

An Aesthetic Alternative

Elliot Eisner (1992, p. 722) has asked "why do we think that all students should be measured by the same yardstick or that we will be able to calibrate the results of different tests in order to make them comparable?" Eisner (1985a) argues that the information provided by achievement testing is not satisfactory. "The measured outcomes that achievement tests provide say nothing about the antecedents of those outcomes ... what we have in achievement test data are consequences, and only a small portion of them at that" (p. 141). He likens achievement scores to the score of a game. The outcome tells nothing about how it was played nor

does it say anything about any problems encountered.

Miller (19--) suggests that evaluation must be growth-oriented and useful in guiding learning instead of the now-prevalent practice of "sorting students" and delivering reward and punishment messages. Sund and Trowbridge (1974, p. 274) say that "to look on a class as a tremendous pool of human potential striving for manifestations is an exciting perception." They also suggest that the function of evaluation should aid in the actualization of that human potential.

Indeed educators, parents, the society in general are encouraging the schools to develop critical thinkers. For too long assessment has been driven exclusively by concerns for measuring and reporting achievement data for outside audiences. Often forgotten, comment Wolf, LeMahieu and Eresh (1992) is the equally important work of internal accountability.

Edmund Short (1990) asserts that educators need a vision of desirable education. If that vision has been lost or was never well stated, then the task, according to Short, is to rethink and reformulate an understanding of what is good and desirable. That vision will involve a look at what is desirable in evaluation and assessment.

Eisner believes that education needs "evaluation methods that exploit the variety of expressive forms through which we understand and make public what we know" (1985, p. 21) Sund and Trowbridge (1974) agree when they emphasize

the importance of a participatory curriculum for each classroom. They reason that:

the structures used in schooling (grade levels, textbooks, standardized tests, and teacher-administration relationships, for example) are created by humans and are amenable to change ... should not be accepted as givens ... but challenged whenever they impede the development of the type of society we want to create in our schools. (p. 11)

Eisner (1985a, p. 165) also protests the climate of a school that places a great deal of emphasis on measured forms of educational performance, but tends to neglect attention to the performer, himself. He notes that a fundamental question that any adequate theory of evaluation should address is not what can be evaluated, or how, or whether or not objectives have been achieved, "but how it is that humans come to know in the first place. And in the second place, how it is that they represent what they know to others" (p. 229).

"Central to looking at children," note Dobson et al. quoting Rogers (cited in Dobson, 1985, p. 9) "is the premise that 'the best vantage point for understanding behavior is from the internal frame of reference of the individual himself'". They go on to note that the child who perceives self as inadequate will generally behave in such manner, as will the child who perceives self as adequate.

Educators are beginning to examine and experiment with

alternative forms of assessment. They are following the lead of many authors and theorists who express the belief that is expressed by Wolf (1988) when he asks:

But what about those questions buzzing in the twilight--the ones about students' abilities to formulate new questions, pursue work over time, arrive at standards of excellence? Those skills simply can't be tapped by highly structured, product-oriented, closed modes of assessment. Information about those skills can only come from looking at student engaged in open-ended, long-term learning where they engage in thinking critically about their own work rather than simply waiting for someone else's "report card". (p. 29)

This then is the challenge to education: to present an alternative that meets the challenges mentioned thus far in this chapter, namely:

1. to begin a movement away from standardized and norm-referenced tests
2. to empower the teacher and provide for reskilling
3. to empower the student, encourage self-assessment and remove labels
4. to address the multiple levels of intelligence found in all children
5. to create a form of assessment that matches a more humanistic language that we use to talk about educating children

Elliot Eisner (1985a, p. 88-91) notes that while

scientific and technological approaches have made some important contributions, they also have several consequences. First they tend to search for generalizations through a process of reductionism. Quality becomes identified with quantity. Second, the technological orientation toward objectives tends to focus on some future state and ignore the present. Third, knowledge itself and children are objectified and reduced to the quantifiable. Fourth, standardized tests standardized the goals and, in fact become the goals. He comments that he believes what is needed in education is not "to seek recipes to control and measure practice, but rather to enhance whatever artistry the teacher can achieve. The new non-scientific approach to educational evaluation suggested by Eisner (1985a, p. 103) is a supplement to current practice that he refers to as educational connoisseurship and educational criticism.

"Connoisseurship," as Eisner (1985a) uses the term, "relates to any form of expertise in any area of human endeavor and is as germane to the problems involved in purse snatching as it is in the appreciation of fine needlepoint" (p. 118).

"Criticism," according to Eisner's (1985a) definition, "is conceived as a generic process aimed at revealing the characteristics and qualities that constitute any human product. Its major aim is to enable individuals to recognize qualities and characteristics of a work or event that might have gone unnoticed and therefore unappreciated"

(p. 118).

Eisner concedes that qualitative forms of inquiry and evaluation are not panaceas (Eisner, 1985a, p. 144). He acknowledges that:

Their methods are demanding, the time it takes to use them exceptionally long, the questions of generalizability difficult, and the verification of their conclusions complex. Yet, because they do provide another view, because they do provide another peak upon which to stand, they promise a great deal.

(p. 144)

This chapter then, has first looked at what are considered desirable features of evaluation and noted the heavy reliance on quantifiable results. Second, it has looked at existing assessment practices and at the advocacy and criticism of these practices and last, it has looked at suggestions for an aesthetic alternative that has as its value base Elliot Eisner's educational connoisseurship and educational criticism.

The following chapter will look at portfolio assessment as an alternative to meeting the criteria presented in this chapter as challenges and as an alternative to meeting Eisner's requirement that "communication about what is happening in schools be part of an expressive, sensitive, and revealing picture of educational practice and its consequences" (Eisner, 1985b, p. 213)

CHAPTER IV

PORTFOLIOS AS AN ALTERNATIVE
ASSESSMENT FORM

The word "assessment", notes David Satterly (1989, p. 1), is from the Latin assidere, "to sit beside". "Sitting beside children suggests a close relationship and a sharing of experience." Satterly notes that it is ironic that educational assessment has come to be associated with one of two contrasting meanings. First, assessment is currently considered to be "hardnosed objectivity, an obsession with the measurement of performances and an increasingly technical vocabulary..." (p. 1). Second, he challenges that assessment is considered a "means by which schools and teachers - wittingly or unwittingly - sort out children for occupations of different status and remuneration in a hierarchically ordered society" (p.1).

In the last chapter it was noted that Eisner and others proposed a new way of looking at the assessment process - one that borrowed its metaphors from the aesthetic paradigm. This idea of educational connoisseurship and educational criticism involves knowing how and what to see and in being able to help others better see, understand and appraise the quality of, in this case, classroom practice. Eisner

(1985a, p. 81) discusses criticism, when applied to art, as the use of methods to heighten perception of the qualities in a work of art. Eisner agreed with Dewey (cited in Eisner, 1985, p. 81), who observed that "the end of criticism is the re-education of the perception of the work of art".

The major virtue of Eisner's concept is that it offers more opportunities to understand what is happening in the classroom and opens up new possibilities for evaluation. One of those new possibilities is found in the use of portfolio assessment.

The purpose of this chapter is to delineate some ways that portfolios might address many of the challenges listed in the previous chapter. Those challenges included moving away from standardized, norm-referenced, summative testing; empowering teachers and students, addressing children's multiple levels of intelligence, creating forms of assessment to match a more humanistic language, encouraging self-assessment and growth, and removing labels.

Portfolios

In their book, Portfolio Assessment in the Reading-Writing Classroom, Tierney, Carter and Desai (1991, pp. vii-viii) list five basic beliefs that are at the root of the ideas in their book. These beliefs grew out of the idea that portfolios could develop classroom practices and traditions reflecting student-centered approaches to

assessment.

1. Teachers are capable professionals who have the capacity to facilitate growth in students when given autonomy and respect.
2. Students are learning how to think for themselves and will work to their greatest capacity and in a creative fashion when given ownership in their learning.
3. Reading and writing are essential survival skills and can lead to a life-long educational process, self-expression, and socio-economic, political and personal empowerment. With reading and writing students can learn to communicate and use ideas effectively.
4. Diversity is inevitable and desirable and the process of education should reflect a diversity of human experience and creative capacity. Defining and standardizing student capacity limits it.
5. The key word in the student-teacher relationship is respect, and it must be mutual and characterized by the understanding that all human beings are worthy and that which they create is worthwhile.

It is the aim of the next section to introduce portfolios, to describe the process of self-assessment, to address the impact of portfolios on students, teachers and parents, and to assess their impact, both positively and negatively.

Portfolios in the Classroom

Why Do We Need Portfolios? Developing artists rely on portfolios to demonstrate their skills and achievements. Within the portfolio they include examples of their work that exemplify the depth and breadth of their expertise. The work they include in their portfolio might indicate their versatility and ability to handle various media. The portfolio might also include several pieces of work on the same subject to indicate their skill. Additionally, one might find work collected over time to demonstrate their growth. Critics and teachers and the artists themselves are better able, with these samples of work, to understand the development that has taken place and to plan for future growth in each area (Valencia, 1990, p. 338).

Valencia (1990) also describes four reasons for a portfolio approach to assessment. First, she notes that sound assessment is anchored in authenticity. Portfolio assessments resemble actual tasks (reading, writing, math, and so on) that a student is required to perform. Second, assessment should be continuous and must show development. This is the difference in simply assessing the outcome (product) and the process of learning over time. Third, assessment should be multidimensional-addressing the many facets of a student's learning. Fourth, assessment should provide for collaboration and reflection by both teacher and student.

As Valencia (1990, p. 338) points out, traditionally

assessment has been done by someone to someone else. Seldom is it viewed as a process that helps an individual evaluate themselves - how well they have learned and what they need to learn next.

The Reading/Language in Secondary Schools Subcommittee of the International Reading Association (1990) argues that informal assessment is an important bridge between formal assessment and classroom instruction. They note that formal assessment provides information about overall achievement and a criteria for deciding if a student is on schedule in relation to other students, but it does not provide the immediate feedback that is necessary for academic growth.

They instead endorse informal assessment because it is "directly related to the curriculum, it provides a meaningful picture of student growth and established attainable goals, helping students maintain a positive self-image, keeping them motivated, and giving them a feeling of purpose and control" (p. 644).

Informal assessment has a personal focus and students use what is being assessed as a focus for continued learning. It can take many forms and generally includes: student self-assessment, teacher observation, teacher designed instruments that monitor development, and projects and assignments selected by the student to represent his/her best efforts. One way to keep an informal record of the student's progress is to use a portfolio, a collection of student work that can show both affective and cognitive

growth. Portfolios are, in a sense, an "expanded definition of assessment in which a wide variety of indicators of learning are gathered across many situations before, during and after instruction" (Valencia, 1990, p. 340).

The terms "performance assessment" and "authentic assessment" are often used interchangeably to refer to the many activities that are a part of informal assessment. They are not, as Meyer (1992, p.39) points out, synonymous. Performance assessment occurs when students are asked to perform specific behaviors that are to be assessed. For instance, students are asked to produce a writing sample to prove they can write. In order to be considered an authentic assessment, the performance is assessed in a context more like that encountered in real life. The performance is not in a contrived context like that of a standardized writing assessment which may involve much structure, limited time, stated topics, required length, and a great deal of teacher direction. Portfolios offer the opportunity to provide authentic assessment.

What Is a Portfolio?. A portfolio, according to Morton (1991, p. 1), is a system for organizing evidence of the literacy development of individual students. Physically, it can be a file folder, an envelope folder, and accordion file, or whatever a person chooses.

Portfolios can be classified according to three types according to Miller (19 --, p. 1):

Type 1 portfolios include singular works to demonstrate

completed mastery or competence (e.g., one of each of four types of writing to show competence in each area). Type 2 portfolios include multiple examples of the same type of product to demonstrate growth on developmental outcomes (e.g., four short stories written over a period of time to document growth). Type 3 portfolios, also called "process-folios", include a piece of work at various stages of completion to document the process used in completion (e.g., note cards, lists of references, outlines, first and final drafts).

What Should Portfolios Contain? There are no rigid rules about what or how much to include in a portfolio. Much of this depends on whether the portfolio is a Type 1, Type 2, or Type 3 as described above. However, Valencia (1990) advises teachers to be selective and include those things which reveal instructional goals. Other literature on the subject of portfolios reveals that the typical portfolio might include (but not necessarily be limited to) any combination of the following:

teacher observation notes, check lists of literacy behaviors, journal entries, writing samples, tape recorded reading samples, running records, conference notes, response logs, reading logs, attitude and interest surveys, interviews, sample of student's best work, written and oral retellings, student self assessments, teacher evaluation and summaries, checklists, brainstorming notes, unit projects, group

projects, notes, outlines, work in progress and at various stages of completion, ideas for new projects, and so on.

The key according to Valencia (1990, p. 339) is "to ensure a variety of types of indicators of learning so that teachers, parents, students and administrators can build a complete picture of the student's development." Portfolios can, as noted by Morton (1991), include any authentic piece of evidence which documents literacy development. The portfolio should not, however, include tests, workbook pages, or other daily graded assignments.

How Much Should The Portfolio Contain? Again, there is no established number of items, but Valencia (1990) advises that the more measures one has, the more reliable will be the conclusions about the child's literacy development. Wolf (1989, 37) observes that students collect a biography of works which "reveal the geology of different moments that underlies the production of any major project." This might include notes, diagrams, drafts and final version of an essay, for example. In addition the portfolio will include a range of works that is deliberately diverse. Reflections about the work are also included in the collection. These things coupled with teacher and student commentaries become the final portfolio that documents the child's growth.

Grant Wiggins (cited in Brandt, 1992, p. 36) contends that just because a task is authentic, it does not mean that it is valid for indicating mastery. Technical people talk

of "generalizability" - does the particular task "generalize" to other similar kinds of tasks? Some people suggest that students may need to do at least six different tasks of a similar kind to make sure inferences are correct about mastery. He notes that it is reasonable to expect a number of samples to be included in a portfolio in order to assess typical performance.

What Evidence Will the Portfolio Contain?. The portfolio should contain materials that allow for the students and teachers to plan together for instruction and literacy activities according to Valeri-Gold, Olson, & Deming (1992). In addition, they consider three additional areas of concern that should be understood:

(a) the focus of the portfolio, based on the clearly defined objectives developed collaboratively by the classroom teacher and the student; (b) the audience, in particular, who examines, critiques, and evaluates the products and processes in the portfolio (e.g., administrators, other classroom teachers, peers); and (3) [sic] the evidence, the pieces of work the student chooses for the portfolio (e.g., works in progress, rough drafts, final copies, checklists, questionnaires). (pp. 299,300)

Since portfolios are supposed to represent real learning activities going on in the classroom, then samples of work from the variety of daily and weekly tasks should be included according to Vavrus (1990):

Literacy process portfolios might include samples of writing folders, excerpts from journals, and projects. A science portfolio might include lab reports, project work, experiments and questions the student has. In math, the portfolio might document improvement in solving increasingly difficult problems, explanations of mathematical processes, or solutions to open ended questions and so on. (p. 52)

How Much Time Is Required for Portfolios? One caution made by Morton (1991) is that one doesn't just "find time" for portfolios in the classroom; it just becomes a routine. As the evidence becomes available to meet the criteria for items to be included, they are placed in the portfolio. Relatively little time is required for actually keeping the portfolio. Some time is required to periodically evaluate the contents and to record the evaluation. With practice this takes no more time than do traditional grading practices.

How and When Will the Classroom Teacher Assess a Student's Work in the Portfolio? As demonstrated by Valeri-Gold, et al. (1992), a timeline can be established for the assessment process. It should indicate if the portfolios will be evaluated in the middle and/or end of a quarter. This timeline might also indicate when conferences will occur and what, if any, anecdotal records will be kept. The selection of works, they note, require the student to be

actively involved in the decision-making process of self-assessment along with the classroom teacher. "Students need to reflect, to monitor their own progress, and to respond to their own learning" (p. 300).

There are numerous check lists available from various sources to help with the process of portfolio evaluation, however, the best lists are those made by the teachers and students as they decide on the criteria for the assessment process that is best for them.

What Happens to the Portfolio at the End of a Term?

There are a number of questions to answer concerning the outcome of the portfolio at the end of the year and, Valeri-Gold et al. (1992) explain, they should be addressed by the student and the teacher as part of the process of portfolio assessment. Will the portfolio be put in a cumulative folder? Will it be returned to the student? Will any or all of the portfolio be kept for the next course of instruction? Valencia (1990) suggests that the teacher and student decide at the end of the year which pieces will remain in the portfolio for the next year and which pieces are ready to go home with the student. This helps communicate the ownership of the portfolio to the student so he or she maintains a personal connection to the work.

How are grades and evaluations communicated?" First, Morton (1991) suggests that teachers communicate their intent before they ever start portfolio assessment.

Although she warns that some parents and administrators may be skeptical, most teachers report that once the portfolios are shown to others, their value becomes obvious. Lapp and Flood (cited in Morton, 1991, p. 6) also suggest including the use of graphs and other visuals which demonstrate growth in specific areas. Reporting progress to parents by way of a narrative report and sharing the portfolio in a conference is much more meaningful than a letter grade on a report card that goes home every nine weeks.

Another way to evaluate the portfolio is for the teacher to tape his/her responses to the portfolio as it is being reviewed. This tape is then sent home to the parent along with the portfolio and the parent is encouraged to sit down with the student and listen to the tape as they examine and read the contents of the portfolio. The tape can then be used several times and can follow the student and the portfolio from year to year.

The Question of Grades

A much harder question to address is that of grades. Grades were designed to accommodate a skills-based definition of learning. As such, a major area of disagreement in assessment practices is that which concerns grades - both letter and numerical. Grades evolved, according to Hargis (1990, p. 12) as the number of students in school increased. Until the mid-nineteenth century when education became more common and widespread, students

received examinations, but there were no grades given. Examinations were not for grades, but to show student progress and to see what areas needed additional work. In learning a skilled trade, for example, an apprentice was judged competent by his master and was then permitted to become a journeyman.

Prior to the introduction of letter and numerical grades, a student's progress was attributable to the teacher. In fact, the teacher was graded based on the performance of the student. Under the subsequent grading systems, the responsibility for achievement was shifted to the student. Grading became the responsibility of the teacher and learning was the responsibility of the student. Achievement became only an index by which the individual could be compared to the standard of his or her grade placement within the school (Hargis, 1990, p. 13).

Out of these grading practices Hargis charts the evolution of the various scales and standards and other schemes to assign grades and scores based on ability levels. In fact, the need to determine these different ability levels led to the development of the intelligence scale by Binet-Simon. Ultimately these standardized intelligence tests developed into the standardized achievement test used today to address performance standards (Hargis, 1990, p. 18).

The current grading practices indicate a powerful need to quantify and this quantification by grades has become an

institutional part of American education. Quantification in the form of grade level divides the curriculum and grouping is based on chronological ages. Students are then ranked by grades based on performance in that particular group.

The curriculum itself is organized around grades. The steps are defined by testing and grading periods, according to Hargis (1990), with all students moving through in lock-step fashion. "...a student's performance is gauged against the normative measure of the difficulty of the material assigned to each grade and step. The lock-step curriculum, in fact, is largely responsible for having grades" (p, 37).

One reason for having grading systems is for teachers to be able to differentiate among students of various academic abilities. The system helps in fitting students into the grades and tracks that are increasingly used as schools become larger and more grade levels and subjects are included in the curriculum. Grades and the leveling that results from them allow students to be segregated into more homogeneous ability groups with the intent of making mass education more efficient (Hargis, 1990, p. 12). Scores' and grades' function, then, seems to be for classifying, grouping, or qualifying students.

Hargis criticizes the grading system as a method which legitimizes giving failing grades:

the only way we can give enough students failing grades is by giving them work that they will fail at doing.

The simplest way to do this is to provide only grade-

level instruction. The grades, themselves, thereby legitimize both failing grades and single level instruction (p. 7).

A number of myths exist concerning grades according to Hargis. Grades are said to be motivating and are said to be objective. In addition, grades are said to be a necessary evaluation tool. Hargis disagrees and first, contends that the only students who are motivated by grades are those who are already getting good grades. "Success is fundamental to achievement," he argues, "and lack of success means lack of achievement." Failing grades are indicative of the failure to provide success for most students who receive them.

Second, grading practices are not objective according to Hargis. Grades are not reliable and, thus, are not completely objective. An "A" in a wealthy suburban area means something different than an "A" in a poor inner-city school. The fact that most colleges do not rely on grades alone, but on SAT and ACT scores, indicates that there is a difference in a letter grade from one school to the next.

Third, letter grades are not particularly useful for evaluation. They say nothing of strengths, weaknesses, readiness, or achievement. Useful evaluation shows what has been learned and what has not and grades, themselves, do not serve a useful instructional purpose (Hargis, 1990).

Glasser (1969) disputes the claim that grades raise academic standards. In fact, he maintains that just the opposite is true. It is his contention that "when grades

become the substitute for learning, and when they become more important than what is learned, they tend to lower academic standards" (p. 65). Poor grades may be used for motivation only if the student is capable of doing better, notes Hargis (1990). Poor grades will not be motivating to the student who is not capable of more. Instead, he or she will be demoralized. As Glasser points out (cited in Hargis, 1990) "all you learn from failing is how to fail" (p. 24).

Letter grades, then, are simply a reflection of the differences in academic skill and achievement that exist in every classroom, given one level of instruction. Glasser (cited in Hargis, 1990) deduces that about 12 percent or three to four in every thirty students will actually fail. This deduction is based on Cureton's observation (cited in Hargis, 1990) that performance relative to grade placement is based on percentage scores which, in turn, are based on curves introduced over 60 years ago.

Portfolio assessment does not lend itself to grades per se. The purpose of a portfolio is to provide an ongoing, meaningful, co-assessment process for a total picture of a student's achievement and to provide a framework for the student's continuous growth and development. The quantification of that process by way of a letter or number has no place in portfolio assessment. If, however, there is a requirement for letter grades, there are a number of creative ways to develop learning contracts or to assign a

point system by which the most valued activities for the portfolio receive the most points. Again, the teacher and the student need to address this question as part of their co-assessment of portfolios. They may offer each other feedback concerning the use of grades and the development of such a system of evaluation in the event that letter grades are unavoidable.

Putting Portfolios in Perspective

With the emphasis in recent years on assessment results, as well as increasing concern about the nature of the most used forms of student assessment there has developed a great deal of conflicting evidence that both supports and questions the performance-based assessments, particularly portfolios. These informal assessments involve the performance of tasks that are valued in their own right as opposed to "paper-and-pencil, multiple-choice tests that derive their value primarily as indicators of correlates of other valued performances" (Linn, Baker & Dunbar, 1991, p. 15). They go on to comment:

Unfortunately, indicators are too often confused with goals, just as norms are too often confused with standards. When this happens, the indicator or norm is apt to lose its value. Worse, the processes that may help to fulfill the fundamental goal often become distorted. The greater the gap between the indicator and the goal, the greater the likelihood of distortion,

particularly when too much emphasis is placed on the indicator. (p. 15)

This lack of correspondence between indicator and goal has become a greater concern since traditional achievement tests have taken on increasingly higher stakes. Linn et al. (1991) note that this call for authentic assessment is not new. In fact, it has been standard advice from some measurement specialists for a long time. Lindquist (cited in Linn et al., 1991, p. 15) argued as early as 1951 that "the most important consideration is that the test question require the examinee to do the same things, however complex, that he is required to do in the criterion situations" (emphasis in the original). Lindquist was not the first to express concern about the effects of testing. Holmes (cited in Maddaus & Tan, 1993) a 19th century British inspector observed first hand, in 1911, the negative effects of tying pupil examination results to teacher salaries:

Whenever the outward standard of reality (examination results) has established itself at the expense of the inward, the ease with which worth (or what passes for such) can be measured is ever tending to become in itself the chief, if not the sole, measure of worth. And in proportion as we tend to value the results of education for their measurableness, so we tend to undervalue and at last to ignore those results which are too intrinsically valuable to understand. (p. 74)

Sixty years later, Ralph Tyler, is reported to have

echoed the same message (cited in Maddaus & Tan, 1983, p. 74) when he warned that society conspires to treat scores on important certifying tests as the major end of secondary schooling rather than as a useful but not infallible indicator of student achievement.

Test Content as Curriculum

Educational assessment is in the process of invention according to Herman (1992, p.74). Old models are questioned; new forms are being developed. While there is potential, what is known is relatively small compared to what remains to be discovered. As has been mentioned in Chapter Three, many researchers have found that a call for accountability pressures both teachers and administrators to plan their curricula based on test content. Researchers conclude that time devoted to test content has narrowed the curriculum.

Consultant Bena Kallick told participants in a mini-conference for ASCD (cited in Update, 1991a, p. 5) that schools could merely "tinker" with alternative assessments or use them to "transform the institution. Since, as she explained, assessment does tend to drive what schools teach and how they teach it, alternative assessments can act as a lever to move schools away from the factory-based delivery of facts model to a new paradigm in which students are active learners and questioning thinkers. It was noted that instruction should shift from imparting knowledge to helping

students locate knowledge and coaching their performances.

Grant Wiggins (cited in Update, 1991a, p. 5) speaking at the same conference called assessment the "Trojan horse of school reform" because of its power to reshape what and how schools teach. He also observed that "we think of assessment as something separate from curriculum and instruction," and insisted that educators should "reintegrate assessment into curriculum ... and build curriculum and assessment out of the same tasks" (p. 5). Similarly, John O'Neil (1992b) supports the idea that better tests can drive more appropriate instruction. His observation is based on experience with direct assessment of student writing, which, he notes, has had a longer history than some other performance assessments.

Judging the Quality of Assessment

New understandings of the nature and context of learning have supported a move toward alternative assessment. As has been mentioned, Howard Gardner's recent work with multiple intelligences has opened the door for developing different components of the mind. While traditional concerns about validity and reliability (involving stability and consistency) are still appropriate, Linn et al. (1991) call for additional criteria for judging the quality of assessment:

Consequences. Consequences involve how people respond to the results of assessment and consequences can influence

the validity of the results. What constitutes a portfolio varies from one setting to another. In addition, the extent to which time is influenced by the way portfolios are used is of interest. It might be reasonable to "inquire about whether the breadth of a student's activities will suffer from overemphasis on a few entries" (Linn et al. 1991, p. 17). There is a necessity to appraise the actual use and consequences of assessment. In addition, Shepherd (cited in Linn et al., 1991) notes that results from standardized tests can be corrupted and it can be assumed that the new forms of assessment would not be immune to similar influences.

Fairness. Judgments about the fairness of an assessment also depend on the uses and interpretation made of assessment results. The assessment should address equally the variety of cultural background of those taking the assessment. It is important to be concerned with biases against racial and ethnic minorities along with gender biases.

Shifting to performance based assessment, however, does not automatically mean there will be equality of performance. In fact, according to Linn et al. (1991, p.18), results from the 1988 National Assessment of Educational Progress (NAEP) and calculated by Langer, Applebee, Mullis, & Foertsch indicate that the difference in achievement between Black and White students was basically the same in writing (which was assessed by essays) and

reading (which was assessed by multiple-choice tests). A 1990 study by Fienberg (cited in Linn et al., 1991) also showed that adding a performance section to the California Bar exam in 1984 did not reduce the differences in passing rates between White and minority test takers. The question that still arises is whether or not all students have had the opportunity to learn what is being assessed.

Transfer and Generalizability. The assessment should support accurate generalizations about student capability. The results should be reliable across raters and consistent in meaning across locales. Research on these issues according to Herman (1992) raises perplexing questions about feasibility. A major concern noted by many including Cannell (cited in Linn et al., 1991, p. 18) is whether standardized tests provide an accurate picture of student achievement or mislead because the scores may be inflated by teaching to the test. It is important to have actual evidence showing that the skills and knowledge that lead to successful performance on an assessment transfers to other tasks.

Cognitive Complexity. It is difficult to tell from looking at an assessment whether it actually assesses higher-level thinking. Many critics argue that standardized tests place too much emphasis on factual knowledge rather than higher order thinking skills. One promise of performance based assessment is that it can place greater

emphasis on problem solving, comprehension, critical thinking and reasoning. But it must also be noted that a hands-on task does not necessarily encourage the use of more sophisticated or complex skills.

The report of the National Academy of Education (cited in Linn et al., 1991) provided a caution:

It is all too easy to think of higher-order skills as involving only difficult subject matter as, for example, learning calculus. Yet one can memorize the formulas for derivatives just as easily as those for computing areas of various geometric shapes, while remaining equally confused about the overall goals of both activities. (p.19)

Schoenfeld (cited in Herman, 1992) refers to the example of an award winning teacher whose reputation was based on his students' Regents Exam. It was found that he had drilled his students on geometry proofs likely to appear on the exam. It was, therefore, difficult to tell if the high scores were due to the acquisition of higher-level thinking or from memorization of algorithms.

Content Quality. The tasks selected to measure a particular content should be worthy of the time and effort for both the test taker and the test rater. Content should reflect important aspects of the discipline. One aspect of the quality of the work is that of the length of time needed to complete it. Simmons reported in a study published in 1990 that his research had led him to guess that students

would work longer on pieces as the year progressed. This was based on a 1976 study by Getzels and Csikszentmihalyi (cited in Simmons, 1990, p. 264) who found that "artists who defined their problems soon after starting a work produced drawings that were less original than those who kept the problem open longer...delay in closure helps to insure that the artist will not settle for superficial or hackneyed problem."

Significantly, Simmons had noted in earlier studies that those who worked the shortest amount of time on papers were generally found to be the most average writers, thus leading him to guess that as the students became better at writing and learned more about their abilities as writers, they would work longer. The data he collected proved otherwise. The further the year progressed, the shorter the duration of writing. It was unclear whether the students worked faster as the year progressed out of increased ability, or whether the papers needed to "sit" longer before students could determine if they were finished. As the time lessened before the end of the year, there was less time to allow a paper to "sit" before final revision. Simmons study did not indicate the relative scoring of the later papers in comparison to those taking longer at the beginning of the year.

One answer to the decreased time spent on work might be that as students become more comfortable with the writing process and more aware of their capabilities, they become

more adventurous and creative and are ready to move on more quickly. This is similar to the way a child dawdles over food he or she is unsure of, but eats quickly that which he or she likes in anticipation of more. Whatever the reason, the fact remains that the content needs to be consistent with the best current understanding of the field and reflect the best aspects of the field. One obvious strategy to assure the content quality of newer assessments is to involve teachers themselves in the review and design of the tasks, materials, and criteria for scoring.

Content Coverage. If there are gaps in the coverage of the content, teachers and students may underemphasize those topics and concepts which are excluded from the assessment. This raises the question of whether assessment tasks represent a full curriculum. The lack of adequate content coverage can lead to misleading scores and distortion of instruction.

Meaningfulness. One rationale for more performance based and authentic assessments like portfolios is that they will encourage worthwhile educational experiences and provide greater motivation for performance. It remains to be seen whether there is a relationship between alternative assessment and student motivation to do well on them. Earlier reference to Hargis (1990) would seem to indicate that those who do well would be positively motivated, while those who do poorly would remain negatively motivated. The

fact that students are able to perform at their personally best instructional level through the use of portfolios would indicate that portfolios could indeed provide a level of meaningfulness for the students.

Cost/Efficiency. Performance-based and authentic assessments such as portfolios are more labor intensive. There is a need to design efficient and cost-effective data collection designs and scoring procedures. Wiggins (1990) notes that not only is cost a major consideration, but so too, is time. Inez Bosworth (cited in Simmons, 1990) comments that "big bucks are involved in portfolios" (p. 262).

This refers to not only the time, space and cost of implementation, but it also can refer to the profit motive that leads publishing firms to develop new and popular curriculum materials that are, in large part, responsible for deskilling the classroom teacher. As with most changes in curricula, the publishers soon develop "packages" to meet the perceived "needs" of the schools and the "needs" of the teachers who are often too limited in time and resources to implement the new alternatives. Published materials can take the locus of control out of the classroom, out of the teachers' hands, and most certainly out of the students' hands. More often than not, these packages actually only meet the need of the publishers to sell more books and increase their profit.

Experts say that performance assessments are likely to

be at least two or three times more expensive per student. This is coming , according to Pamela Aschbacher in her 1990 survey for the Center for Research on Evaluation, Standards, and Student Testing (CRESST) at University of California-Los Angeles (cited in Herman, 1992), as "we're at a time when a lot of people are cutting their budgets" (p. 18).

In a year already crowded with other tests and activities, teachers are often not able to meet the demands of the extra class time and preparation time needed for new assessment forms like the portfolio. O'Neil (1992a) notes that the time spent on administering the performance itself "pales in comparison to the time needed for teachers to revamp their instruction to better prepare students for the new tests" (p.18). Both the class time and the preparation time are higher for the teachers who aim to meet the demands of the the new assessments. Because pencil and paper tests are efficient and comparatively inexpensive many school systems will continue to use these instead of exploring the use of portfolios and other alternative assessments.

Other Concerns

Although these areas of concern address a number of the problems encountered in adopting alternative assessment procedures, there are a few other problems. First, there appears to be little agreement on how to demonstrate proficiency in certain tasks. Teachers, who have long been conditioned to remain objective, often have a difficult time

learning to trust their own ability to judge. They are often encouraged and warned to be able to justify their grades or evaluations with percentages, scores, and so forth - an indication that their professional judgment is of little or no value.

Second, there is often little continuity from one task to another. The use of portfolios is often isolated in one area or discipline. For instance, the students will use this approach in the language arts, but not in social studies or math. In addition the tasks are often fragmented with some of the assessments being given for the sake of assessment and bearing little relation to any other activity or task.

Third, there seems to be little at this point to ensure the integrity of the standards or to maintain interrater reliability. There is often not enough objectivity according to some. There are few guidelines or models of achievement. Indeed, there are some who suggest that to develop guidelines is to invite over-standardization.

Fourth, there is difficulty in assessing and evaluating such areas as effective problem solving and cooperative collaborating. How does one evaluate contribution to the community? What kind of scale or scoring rubric is used to communicate achievement in those areas - if achievement in such areas can even be defined?

Assessment As an Occasion for Learning

Wolf, Bixby, Glenn and Gardner (1991, p. 57) write that assessment should be an occasion for learning. Portfolio based assessment can provide a context in which students regard the assessment as a learning process. Grant Wiggins comments that "performance assessment is for probing and prodding the student's mind to reveal what it knows and can do in action" (cited in Willis, 1990, p. 4). According to Wolf (1989) when students maintain portfolios of their work, they learn to assess their own progress as learners, and teachers gain new views of their accomplishments in teaching. The result of a portfolio experience is not a score on an exam, but instead, a reflection on a sample of work.

In the beginning of this chapter there were mentioned several areas of concern that would eventually be addressed in this paper. Portfolio assessment can and does offer some solutions to the problems often encountered when dealing with standardized curricula and testing programs. The following discussion recognizes some of these solutions.

The Move Away From Standardized Tests. Wolf, Mahieu and Ersh (1992) contend that the portfolio replaces what used to be the quarterly test, but ideally it does more. Not only does it allow students to think about what they have learned and what they still need to learn, but also it allows for reflection of worth. In olden times the

competence of a person was judged by the quality of his/her work-output, the material used, the attention to detail and creativity...not a test. An analogy is the question of whether or not to judge a carpenter by a paper-and-pencil exam or by samples of an actual building project.

Similarly, there is a difference between being able to identify brush styles and sizes, between being able to name and discuss the differences in paint pigments, between being able to define various canvas preparations and in being able to paint a masterpiece as did Van Gogh.

The Need for Empowerment of Teachers and Students.

Frazier and Paulson (1992) discuss portfolios as an asset: portfolios offer students a way to take charge of their learning. They also encourage pride, ownership, and high self-esteem. Wiggins (1990) comments that teachers need to re-think roles to become coaches or "enablers". O'Neil (1992a) observes that portfolios have the potential to draw teachers back into the heart of assessment. The teacher can be reskilled and can become "empowered" Herbert (1992) reported that a portfolio project used in her classes had positive effects on self-esteem of the students and the professionalism of teachers and, as a result, a more effective faculty emerged.

Clay affirms that teachers are empowered as they check their own assumptions against what they find out in systematic observations. They can be encouraged by what Johnson (cited in Clay, 1990) refers to as "intelligent

unrest" - to search for explanations of what they find out. This leads to teachers becoming active constructors of their own program and evaluations and allows them to, as Frazier & Paulson (1992) exclaimed, suddenly work together. They note one teacher's comment that: "I had become their collaborator instead of their teacher" (p. 62).

The Recognition of Multiple Levels of Intelligence.

"Children who do not learn to see will not be able to write, not because they cannot spell, but because they will have nothing to say" (Eisner, 1985, p. 175). Sullivan's split-brain literature (cited in Arnowitz and Giroux, 1985) argues that cognitive and creative sides of the brain be given equal attention. "This 'heurism' shows that the brain is more than a mere computer. It has artistic features and should not be limited to quantitative abstraction" (p.18). Gardner (cited in Brandt, 1988) of course has done extensive study to develop his theory of multiple intelligences and agrees that portfolios can provide a method to allow the development of the different aspects of a child's intelligence.

The Need For a More Humanistic Language. The

prevailing language of the factory model has included the use of terms commonly associated with an assembly line. Students are treated like products. Classroom teachers are called managers. There is output to be measured and quality control in terms of cost effectiveness and efficiency.

Students who don't succeed are weeded out as defective. The cynical, yet true rubric often quoted by the manufacturing world to describe quality control is "beat to match, bend to fit, weld to hold and paint to cover." Unfortunately, that is often how students are treated by structured school systems that seek to grade, label and standardize.

Wolf et al. (1991) suggest that like other "performers" someone engaged in portfolios sustains a "long arc of work...thinking involves rehearsals, revisions, criticisms, and new attempts arranged in nothing like the straightforward ordering of the scientific method" (p.34). The factory model looks at mastery - indicating no room for growth beyond what is mastered. Portfolios look at continual development and improvement rather than at mastery. "Errors aren't something to be tallied, notes Hargis (1990), they are to provide corrective feedback" (p. 61). A technically correct paper may lack life while a lively, interesting and creative paper may lose emphasis and impact because of poor grammar, spelling and punctuation.

The Need For Self-Assessment. Students should be able to regard portfolios and the assessment that accompanies them as an occasion for learning. As a result of portfolios, teachers and students have access to a continuous body of work in which to discover characteristic patterns, style, and difficulties needing additional work. (Wolf et al., 1991, p. 58)

Students should be encouraged to engage in self-

reflection according to Frazier and Paulson (1992). With self-reflection that is an outgrowth of portfolio assessment Eisner (1985) explains, students will develop unique meanings and will make their own outcomes. Wiggins (1990) makes the point that:

the whole point is to put the student in a self-disciplined, self-regulating, self-assessing position. Portfolios can help in this process, but it's going to be stymied if we're still defining student performance in terms of control over bits of information. (p. 51)

The Need to Do Away With Labels. Simmons (1990, p. 265) challenges the educational community to "stop underestimating student ability, be more fair to our weakest writers, and profile both the habits and judgments of student writers..." through the use of portfolios. Clay (1990) admits that test scores have been useful for classification and comparison of students but also holds that classifying and comparing are several steps removed from teaching them effectively. Low student outcome scores do not distinguish between three things which, according to Clay, might be responsible for the lower scores: "low prior achievement of students, or poor program delivery, or ineffective education policies" (p. 289).

The Benefits of Portfolios

Portfolios according to Tierney et al. (1991) are not objects. They are vehicles for ongoing assessment by

students. They represent activities and processes (selecting, comparing, self-evaluation, sharing, goal setting) more than they do products.

Portfolios offer a new framework for assessment- one that facilitates student reflection in conjunction with reading and writing- a framework that respond to demands for student empowerment, the changing nature of classrooms, and a new consensus regarding the need for revamping testing practices. Portfolios offer a framework that is dynamic and grounded in what students are actually doing. (p. 42)

The benefits of portfolios are probably most apparent when compared to what seems to be the traditional practices. Tierney (1991) summarizes seven differences in assessment processes and outcomes between portfolios and standardized testing practices.

First, portfolios represent the range of reading and writing students are engaged in while testing assesses students across a limited range of reading and writing assignments which may not match what student do. Second, portfolios engage students in assessing their progress and/or accomplishments and establishing on-going learning goals while testing is mechanically scored or scored by teachers who have little input. Third, portfolios measure each students achievement while allowing for individual differences between students but testing assesses all student on the same dimensions. Fourth, portfolios

represent a collaborative approach to assessment while the testing process is not a collaborative one. Fifth, portfolios have a goal of student self-assessment but student assessment is not a goal of testing. Sixth, portfolios address improvement, effort and achievement while testing addresses achievement only. Seventh, portfolios link assessment and teaching to learning while testing separates learning, testing and teaching.

This chapter has sought to discuss one alternative assessment procedure known as the portfolio. It has included some of the commonly asked questions about portfolios and their use. In addition it has addressed some concerns being voiced about the inadequacies of standardized tests to evaluate educational progress. As Elliot Eisner (1993) has noted in discussing the difference between standards and criteria:

criteria facilitate the search for qualities we value within an essay, a scientific experiment, a painting, a work of history and the like. These works, Dewey argues and I concur, are not susceptible to measurement by standards, although they are amenable to appraisal by criteria. (p. 22)

The aesthetic paradigm set forth by Eisner (1985) has as its basis the idea of educational connoisseurship and educational criticism. The connoisseurship (or appreciation) is a private event. The externalization of what one appreciates must also occur and it is here that Eisner

encourages the aim of criticism. Its aim is:

to re-educate one's perception of the phenomena to which one attends. Criticism is a written or spoken statement about something. Educational criticism is about educationally important matters. Because criticism is in large measure-although not exclusively-an artistic creation (the critic must render the qualities he has perceived in a form capable of eliciting, in part, an empathetic understanding), the skills the critic needs are significantly artistic in character. (p. 380)

As Eisner also notes, the time has come to look at other views and if the artistic view is a promise for a new perspective, then the portfolio is a way to approach this critical aspect of education.

The following chapter will present conclusions, implications and recommendations that have come as a result of this study and analysis.

CHAPTER V
IMPLICATIONS, CONCLUSIONS,
AND RECOMMENDATIONS

Introduction

The previous chapters were devoted to fulfilling the basic purpose of this study - a critical analysis of Portfolios as an alternative assessment tool. Chapter II was devoted to examining the foundations of the dominant paradigm, that of the Factory Metaphor. Chapter III was devoted to examining the concerns, criticisms and defenses of assessment, including a discussion of existing assessment practices and the suggestion of an aesthetic alternative. Chapter IV was devoted to examining the use of Portfolios as an alternative assessment form influenced by the aesthetic paradigm and included a rationale and perspective concerning their use.

The present chapter completes this study by, first discussing the implications of using the Portfolio method. In addition, the second section of this chapter will discuss some of the conclusions drawn from this study and a third section will offer some recommendations for future studies concerning the use of portfolios.

As Herman (1992) reminds us, "education is in the

process of invention...what we know is relatively small compared to what we have yet to discover" (p. 74).

What we do know, according to Wolf et al. (1992), is that assessment has been driven too exclusively by concerns for the measuring and the reporting of achievement data for outside audiences with little or no internal accountability or thinking about what students should know or how they should learn it. In fact, it has been noted that Apple (1990), severely criticizes current curriculum practices which focus on training and as such train some to lead and others to follow. He further condemns practices that do not encourage students to think for themselves, but instead allow the so-called intellectual leaders to think for them.

Meaningful learning is reflective, constructive and self-regulated according to cognitive researchers Bransford and Vye, Davis and Maher, Marzano, and Wittrock (cited in Herman, 1992). To know something is not just to have received information, but to have interpreted it and related it to other knowledge one already has. The question then, becomes one of how to meaningfully involve students and how this involvement will empower them and help them to internalize standards of quality.

Recent movements to re-define outcomes of worth and facilitate self-directed learning have called for student evaluation that is multi-dimensional, personalized and useful instructionally. Portfolio assessment is becoming a popular and viable alternative to achieve these goals.

Implications

As was noted in Chapter IV, those who support the use of the portfolio as an alternative assessment practice make six claims for its use in the classroom. Their claims include first, that portfolios can help move the curriculum away from standardized testing as the sole evaluation of a student's achievement. Second, they claim that the use of portfolios can help teachers and students become empowered. A third claim is that the use of portfolios can allow for the recognition of the multiple levels of intelligence that are possessed by all students. Fourth, they claim that portfolios can address the need for a more humanistic language in the classroom. Fifth, they note that portfolios can allow for self-assessment and sixth, they point out that portfolio assessment can help do away with labels that are used to classify and sort children.

If portfolios are chosen as an alternative, what will be the significance of their use to guide the field of curriculum evaluation? What, in other words, will be implied by the choice of portfolios? Six implications are discussed below.

First, the acceptance of portfolio assessment implies that cost and efficiency associated with cost are not to be the guiding factors in assessment. As Nuttall (1992) has cautioned, "the cost of portfolio assessment both financially and in terms of time is immense" (p. 57).

A consequence of this implication is that states and

school systems will be forced to address the funding and expense of evaluation, in general. It must be noted that including the assessment as a part of the general classroom instruction rather than as a separate function would possibly prove more effective in terms of time and efficiency.

Second, the acceptance of portfolio assessment implies that assessment is value laden and judgmental. Again Nuttall (1992) notes that performance assessment, and by implication, portfolio assessment, will demand new skills of the teacher as well as a new professionalism. This type of assessment can only take off with the will and expertise of the whole profession and with the trust of the parents and politicians. Rather than apologize for using personal judgment, a teacher should be able to command the same respect for using judgment that, say, a doctor would receive.

A consequence of accepting the judgment of teachers as a natural part of assessment is that the teachers will become more empowered. The effects of deskilling will be neutralized and teachers, once again, will be able to practice the art and craft of teaching, rather than the monitoring and accounting skills which relegate them to reporting norms, grades, scores and so forth.

A third implication of the acceptance of the portfolio as an assessment tool, is the recognition of multiple intelligences and the fact that they need to be addressed by

more than one approach in the classroom. Wolf et al. (1991) insist that education should be multi-dimensional. The recognition of multiple intelligences follows to some extent the recognition of multiple aptitudes that has long been a part of the program for the gifted. If an approach which involves multiple dimensions works for the gifted, then it should also be a profitable approach for all other students. A logical consequence is that standardized tests, which address only one level of intelligence, will cease to become the main or only form of assessment used in the curriculum.

A fourth implication arising from the use of the portfolio is that assessment must be authentic. Students must be evaluated in the same manner as they are expected to use their skills or knowledge. Authentic measures involve performance of tasks that are valued in their own right. In contrast, "paper-and-pencil tests and multiple-choice tests derive their value primarily as indicators of correlates of other valued performances" (Linn et al., 1991, p. 15).

In addressing authentic assessment, Kemp (1985) notes that a learner should anticipate being tested in the same type of behavior as indicated by the objective. Thus the verb in the objective should alert the learner to the emphasis necessary when studying the content (p. 161). "There must be a direct relationship between learning objectives and test items...the verb component of the learning objective indicates the form that a test item should take: 'to solve' should indicate writing a solution"

(p. 161). Lindquist (cited in Linn et al., 1991) agreed that the most important consideration is "that the test questions require the examinee to do the same thing, however complex, that he is required to do in the criterion situations" (p. 15).

A consequence of attention to authentic assessment is that schools will be forced to abandon the exclusive use of purely quantitative methods. Grades, percentages and scores will be rejected in favor of more criterion-referenced, formative, and critical evaluation procedures.

A fifth implication of portfolio assessment is that the student will be considered equally valuable in the assessment process. Miller (19) notes that evaluation must stress self-evaluation, the process of making judgments about the worth of one's work for decision-making. As a consequence, schools will be forced to look at evaluation that is relative, appropriate, and adaptable and that looks at the outcomes of worth that were emphasized during the instruction and learning. Additionally, students must be allowed to become active participants in the evaluation process.

A sixth, and somewhat negative, implication associated with the use of portfolio assessment is that there will be the tendency to attempt new approaches and then quantify the results. Portfolios could simply be reduced to projects that involve percentages, scores, letter grades, and so forth. The consequence is that the results can be corrupted

and the original intent of portfolios will be subverted in an attempt to maintain the current standards of assessment.

Conclusions

This study was introduced by calling attention to the fact that assessment practices are in need of change. The language that is often used in education fields refers to education as democratic, process oriented, child centered and learner directed with the teacher as a facilitator, a pedagogue. However, the methods that are used are more often those which seek to quantify.

The dominant paradigm of the Factory Model seeks to batch process, label, pigeon-hole, and norm students who are treated like products on an assembly line. This model has evolved from the behaviorist and scientific tradition which is structured, managed, and controlled. The assessment practices currently in use are part of the behaviorist objective of accountability, measurement, standardization, and transmission of facts.

In addition to noting the reliance on a system best characterized by the word "control", this study went on to discuss the historical, democratic, religious, and philosophical roots of this dominant paradigm. This study examined both the advocacy position and the critical position of current evaluation and assessment practices. Finally, this study offered the portfolio assessment as an alternative to the dominant use of standardized tests as

evaluation of achievement.

In relation to this study, five conclusions are presented:

First, the dominant paradigm of science and factory metaphors is at odds with how educators say that they view students. The teacher's role in the factory model is that of a manager rather than a leader or facilitator and the student is a learner who is to be taught and tested. The aesthetic paradigm proposed by Eisner, out of which portfolio assessment has developed, focuses on the student and teacher as collaborators. Instead of sorting students and delivering reward and punishment messages, as mentioned by Miller (1990), evaluation must be growth-oriented and useful in guiding learning.

A second conclusion is that traditional evaluation methods need to be expanded to address theoretical concepts of education. An important outcome of the alternative assessment movement is that it challenges the education community at large to reconsider just what are called valid interpretations of any kind of assessment information according to Linn et al. (1991). Eisner (1985a) notes that numbers cannot convey all that needs to be said about the qualities that constitute educational objectives or events. He observes that numbers are reporting devices. Their meaning derives from the scale of which they are a part and the referents they are used to represent. Numerical indices as surrogates are not self-explanatory without knowing their

position on a scale and the qualities they are selected to represent.

In addition, Maddaus (cited in Sears, 1990) documents the effect of testing and concludes that "measurement-driven instruction is, and always has been, devastating to both the curriculum breadth and teaching flexibility needed to ensure high quality education" (p. 202). As Eisner (1985a, p. 225) notes, because objectivity is seen as such a prime virtue in evaluation, qualities that are hard to measure reliably are often altogether neglected in favor of those that are more easily measured. This, in turn, influences the kind of opportunities students will have and the kind of mental skills they are able to develop.

A third conclusion is that current methods that rely on standardized testing developed outside the classroom are part of the process that deskills teachers and fails to give all students a chance to succeed. Under current practices, students are ranked against one another rather than on how well they do what they do. The use of alternative assessment in general, and portfolio assessment in particular, can engage children in interesting work and provide a "test" worth teaching to. While teachers will find the Portfolio somewhat more work than grading true-false tests, they also will likely find that they are once again designing, implementing and evaluating their programs personally instead of implementing a "package".

A fourth conclusion is that the current curriculum is

grade and score driven. Grades are an obstacle to the implementation of many of the alternative assessment systems. Measurement for grades takes the focus of assessment from its more important function as a tool for learning according to Hargis (1990). According to Herman (1992), if teachers have tested the student's ability to think and to reason, then every other area they encounter is available to the student when their need for the knowledge is there, when the knowledge has meaning.

Assessment is a static process because of grades according to Hargis (1990). Assessment should provide helpful immediate feedback. When testing for grades, an error is no more than a red mark on a paper to be tallied. Errors, he cautions, should not be left uncorrected. Errors should provide good opportunities to learn, to discuss, to question, and to think. An error uncorrected or unaddressed is simply an error - nothing more. Hargis (1990) cautions that errors on work are not something merely to be tallied, they are to be used to provide corrective feedback. "Errors uncorrected are errors practiced" (p. 61).

A fifth conclusion is that the scientific and factory based metaphor is not compatible with the aesthetic metaphor as a foundation for the curriculum. This conclusion is drawn from the work of Pereles' 1987 doctoral dissertation concerning conflicting paradigms.

Pereles (1987) writes that "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" (p. 133). In addition, she notes six dimensions in which these two approaches differ. The first difference is the focal point. The scientific mode focuses on observable behavior, while the aesthetic mode is concerned with what is non-observable. The main focus is understanding the meaning of experience.

A second difference is in emphasis. The scientific approach is outcome oriented. It emphasizes measurement and experimentation. The aesthetic approach, on the other hand, is process oriented. Criticism is used to describe and interpret processes in the classroom.

A third difference is in the source of evaluation data. The scientific approach collects data through standardized instruments. The aesthetic approach uses the human instrument (the critic) as the source of data.

A fourth difference is the form of representation that is used. The scientific approach uses a discursive form of representation, i.e. numbers. Data is expressed in quantitative terms. The aesthetic approach uses non-discursive forms of representation; e.g., poetic language. These forms render an account of an event or experience.

A fifth difference is in the ultimate aim of each approach. The scientific approach aims to produce laws that can provide a basis for predicting and controlling outcomes. The aesthetic approach aims to understand. It does not aim

to predict or control.

Finally, a sixth difference is in the criteria used to judge the validity of the evaluation. The scientific approach is concerned with objectivity. The aesthetic approach acknowledges the influence of the evaluator. According to Eisner (cited in Pereles, 1987) the artistic evaluator is neither a neutral observer nor a disinterested interpreter.

As Pereles notes, these six differences between scientific and artistic approaches to curriculum evaluation provide support for the conclusion about their incompatibility.

Recommendations

First, Herman (1992) suggests that changes in assessment are only part of the answer to improved instruction and learning. Schools need to support and implement new instruction strategies and to institute other changes to assure that all students can achieve the complex skills that these new assessment procedures strive to represent.

Colleges, in seeking to be at the forefront of the alternative assessment movement, might provide a portfolio experience to enhance the learning in teacher education classes and to provide a model for later use by education students in their own classrooms. Miller (19) encourages the expansion of current efforts to broaden the ways in

which evaluation occurs. Part of this expansion would be to train teachers in suitable techniques of alternative assessment. Miller also encourages the use of the portfolio as a method for documenting growth, priorities, competence and potential and points out that it can be used as a centerpiece for celebrating the student's growth towards becoming a teacher.

Second, Wiggins (cited in Brandt, 1992a) is critical of the fact that American education has a history of trying to re-invent the wheel. If the business community will continue to be a major influence on education, then one recommendation is to borrow at least the positive aspects of this influence and begin to evaluate as the business community does, using real and authentic tasks.

Third, since most developers of alternative assessment and portfolio programs are at the design and prototype stage, the suggestion is for trials in these various areas to study validity, reliability, and generalizability in terms of portfolio assessment versus standardized assessment. These trials should also seek to study the magnitude of variability due to the raters and to the sampling tasks. This could include studies to indicate whether abilities assessed by portfolios or other performance assessments can be generalized from task to domain, from one day to the next, and from one rater to another.

Fourth, there should be studies of motivation to see

how students respond to the portfolio technique. In addition, these studies should inquire how students actually view portfolio assessment. The studies should also investigate student and teacher understanding of performance based assessment.

Another recommendation is to study teachers' ability or perceived ability to use subjective judgment and to trust that judgment. Part of this study should seek to discover if teachers do subjectively grade and then convert scores to objective scores. In addition, do students feel more judged when the teacher marks a score on their work, when the teacher writes comments, or when the teacher comments verbally? Which format do they prefer?

Fifth, future studies might address the psychological and sociological aspects of the aesthetic paradigm. A discourse analysis of various metaphors should prove interesting in analyzing, not only the sociological and psychological, but also the philosophical impact of metaphors and the paradigms out of which they have developed.

Finally, another question would be whether there should be local, state, or national assessment standards. Considering the current push for national testing and evaluation, an analysis of this issue could be undertaken. Wolf et al. (1991) have already noted the development of what they call a "schizophrenic" approach to educational evaluation in which one form of assessment is used at the

local level and another form of assessment is used at the national level.

What seems to be happening is that many educators are trying to re-invent an educational system that is the parallel of the very economic system Eastern Europe is walking away from - namely a centrally run, centrally designed, centrally mandated "command form of government" (Brandt, 1992, p. 37). It's going to fail in education, he notes, just as it failed in economics because it doesn't empower and energize the entrepreneurship of local people. If the vision was lost or was never well articulated, as Schubert (1986) suggests, then the task is to rethink and reformulate an understanding of what is good and desirable in a good curriculum. One cannot pursue alternative assessment merely on personal conviction that it is better, cautions Wolf et al. (1991).

One needs to develop rigorous standards, rules of evidence, and realistic expectations. New modes of assessment are much more accurately described as rediscovered modes, they note. If this movement is to be more than a "flurry", then educators have to be as tough minded in designing new options as they are in criticizing available options.

The design and implementation of alternative modes of assessment entail nothing less than, as Wolf et al. (1991) call it, "a wholesale transition from what we call a testing culture to an assessment culture" (p. 58). Qualitative

forms of inquiry and, by implication, qualitative forms of assessment such as the Portfolio offer no panaceas for educational problems according to Eisner (1985a):

 Their methods are demanding, the time it takes to use them is exceptionally long, the question of generalizability difficult, and the verification of the conclusions complex. Yet, because they do provide another view, because they do provide another peak upon which to stand, they promise a great deal. (p. 144)

BIBLIOGRAPHY

- Apple, M. W. (1990). Ideology and curriculum. New York: Routledge.
- Arrowitz, S., & Giroux, H. A. (1985). Education under seige: The conservative, liberal and radical debate over schooling. Massachusetts: Bergen & Garvey Publishers, Inc.
- Brandt, R. (1988, January). On assessment in the arts: A conversation with Howard Gardner. Educational Leadership, 45(4), pp. 30-35.
- Brandt, R. (1992, May). On performance assessment: A conversation with Grant Wiggins. Educational Leadership, 49, pp. 35-37.
- Brandt, R. (1992, May). Overview: A fresh focus for curriculum. Educational Leadership, 49, p.7.
- Burke, A. (1991). Portfolios in chapter I evaluations. (Report No. UD 028 478). Portland, OR: Oregon Educational Research Association. (ERIC Document Reproduction Service No. ED 340 806)
- Clay, M. (1990, March). Research currents: What is and what might be in evaluation. Language Arts, 67(3), pp. 288-297.
- Dobson, R., Dobson, J., & Koetting, J.R. (1985) Looking at, talking about, and living with children: Reflections on the process of schooling. Lanham, MD: University Press of America, Inc.
- Eisner, E. W. (1985a). The art of educational evaluation: A personal view. Philadelphia, PA: The Falmer Press.
- Eisner, E. W. (1985b). The educational imagination: On the design and evaluation of school programs. New York: Macmillan Publishing Co.
- Eisner, E. W. (1991). The enlightened eye: Qualitative inquiry and the enhancement of educational practice. New York: Macmillan Publishing Co.

- Eisner, E. W. (1992, May) The federal reform of schools: Looking for the silver bullet. Phi Delta Kappan, 722-723.
- Eisner, E. W. (1993, February). Why standards may not improve schools. Educational Leadership, 50(5), pp. 22-23.
- Ellis, A. K., Cogan, J.J., & Howey, K.R. (1991). Introduction to the foundations of education. Boston, MA: Allyn and Bacon.
- Frazier, D. M., & Paulson, F. L. (1992, May). How portfolios motivate reluctant writers. Educational Leadership, 49, pp. 62-65.
- Glasser, W. (1969). Schools without failure. New York: Harper & Row.
- Goodlad, J.I. (1979) What schools are for. [Monograph]. Phi Delta Kappan Educational Foundation.
- Goodlad, J.I. (1984). A Place Called School. New York: McGraw-Hill Book Co.
- Grant, G. (Ed.). (1991). Review of research in education - 1991. Washington, DC: American Educational Research Association.
- Hargis, C. H. (1990). Grades and grading practices: Obstacles to improving education and to helping at-risk students. Springfield, IL: Charles C. Thomas, Publishers.
- Herbert, E. A. (1992, May). Portfolios invite reflection from students and staff. Educational Leadership, 49, pp. 58-61.
- Herman, J. L. (1992, May). What research tells us about good assessment. Educational Leadership, 49, pp. 74-77.
- Kemp, J. E. (1985). The instructional design process. New York: Harper & Row.
- Koetting, J.R. (1990) Audio-visual communication strategies: CIED 5753. Tulsa, OK: Kinko Professor Printing.
- Kuhn, T. (1971). The structure of scientific revolutions. Chicago: The University of Chicago Press.

- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991, November). Complex performance-based assessment: Expectations and validation criteria. Educational Researcher, 15-21.
- Maddaus, G.F., & Tan, A. G. (1993). The growth of assessment. In G. Cawelti (Ed.), Challenges and achievements in American education: 1993 yearbook (pp. 55-79), Alexandria, VA: Association for Supervision and Curriculum Development.
- Meyer, C. A. (1992, May). What's the difference between authentic and performance assessment? Educational Leadership, 49, pp. 39, 40.
- Miller, M. A. Evaluating with portfolios: Towards improving richness of student work. [Monograph]
- Morton, J. (1991). What teachers want to know. (Report No. CS0 10 710). (ERIC Document Reproduction Service No. ED 336 728)
- Nuttall, D. K. (1992, May). Performance assessment: The message from England. Educational Leadership, 49, pp. 54-57.
- O'Neil, J. (1992a, May). Putting performance assessment to the test. Educational Leadership, 49, pp. 14-19.
- O'Neil, J. (1992b, June). National standards: Assessment system draws mixed response. ASCD Update, 34, p. 7.
- O'Neill, W. F. (1983). Rethinking education: Selected readings in the educational ideologies. Dubuque, IA: Kendall/Hunt Pub. Co.
- Pereles, G.E. (1987). An examination of the dominant and the alternative paradigms in curriculum evaluation. Unpublished doctoral dissertation, Oklahoma State University.
- Reading/Language in the Secondary Schools Subcommittee. (1990, May). Portfolios illuminate the path for dynamic, interactive, readers. Journal of Reading, 644-647.
- Rowntree, D. (1987). Assessing students: How shall we know them?. New York: Nichols Publishing Co.
- Satterly, D. (1989). Assessment in schools. New York: Basil Blackwell.
- Schubert, W. E. (1986). Curriculum: Perspective, paradigm, and possibility. New York: Macmillan Publishing Company.

- Sears, J. T., & Marshall, D. (Eds.). (1990). Teaching, and thinking about curriculum: Critical inquiries. New York: Teachers College Press.
- Selakovich, D. (1967). The schools and American society. Massachusetts: Blaisdell Publishing Company.
- Short, E. C. (Ed.). (1991). Forms of curriculum inquiry. New York: State University of New York.
- Simmons, J. (1990, March). Portfolios as large-scale assessment. Language Arts, 67(3), pp. 262-267.
- Sund, R. E., & Trowbridge, L. W. (1974). Student-centered teaching in the secondary school. Columbus, OH: Charles E. Merrill Publishing Co.
- Szatjn, P. (1992, November). A matter of metaphors: Education as a handmade process. Educational Leadership, 50, pp. 35-37.
- Tierney, R. J., Carter, M.A., & Desai, L.E. (1991). Portfolio assessment in the reading-writing classroom. Massachusetts: Christopher-Gordon Publishers, Inc.
- Tyler, R. W. (1949). Basic principles of curriculum and instruction. Chicago, IL: University of Chicago Press.
- Valencia, S. A. (1990, January). A portfolio approach to classroom reading assessment. The Reading Teacher, 43(4), 338-340.
- Valeri-Gold, M., Olson, J.R., & Deming, M.P. (1992, January). Portfolios: collaborative authentic assessment opportunities for college developmental learners. Journal of Reading, 298-305.
- Vavrus, L. (1990, August). Put portfolios to the test. Instructor, 100(1), pp. 48-53.
- Wiggins, G. (1990, August). A conversation with Grant Wiggins. Instructor, 100(1), p. 51.
- Willis, S. (1990, September). Transforming the test. ASCD Update, 32, pp. 3-6.
- Wolf, D. (1988, January). Opening up assessment. Educational Leadership, 45(4), pp. 24-29.
- Wolf, D. (1989, April). Portfolio assessment: Sampling student work. Educational Leadership, 46(7), pp. 35-39.

- Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1991). To use their minds well: investigating new forms of student assessment. In G. Grant (Ed.), Review of research in education - 1991 (pp. 31-45). Washington, DC: American Educational Research Association.
- Wolf, D., LeMahieu, P. G., Eresh, J. (1992, May). Good measure: Assessment as a tool for educational reform. Educational Leadership, 49, pp. 8-13.

VITA 2

Sandra Douglas Massey

Candidate for the Degree of

Master of Science

Thesis: A CRITICAL ANALYSIS OF THE PORTFOLIO METHOD AS AN
ALTERNATIVE ASSESSMENT FORM

Major Field: Curriculum and Instruction

Biographical:

Personal Data: Born in Tulsa, Oklahoma, January 23, 1947, the daughter of E.L. and Maxine Douglas. Married Lee R. Massey, May 29, 1969 in Tulsa, Oklahoma. Two children, Jennifer LeeAnn Massey and Scott Douglas Massey.

Education: Graduated from Daniel Webster High School, Tulsa, Oklahoma, in May, 1965; received Bachelor of Science Degree in Language Arts, Secondary Education from Oklahoma State University in August, 1969; completed requirements for the Master of Science Degree at Oklahoma State University in July, 1993.

Professional Experience: Speech Director and English Teacher at LDF Community Schools in LeGrand, Iowa from September 1970 to May, 1973. English Teacher for Union Independent Schools in Tulsa Oklahoma from January, 1974 to May, 1975. English Teacher and Administrative Assistant for Wakefield Academy in Tulsa, Oklahoma from September, 1981 to present. Principal at Wakefield Academy from January, 1991 to present.

Professional Organizations: Association for Supervision and Curriculum Development.