INFORMATION TO USERS

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

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

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

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

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6” x 9” black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700  800/521-0600
A CASE STUDY OF THE INTRODUCTION OF
AUTHENTIC ASSESSMENT TO A DISTRICT,
ITS TEACHERS, AND A SENIOR CLASS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Sandra K. Paxton, M.A.

The Ohio State University
1996

Dissertation Committee:
Professor Suzanne K. Damarin, Advisor
Professor William E. Loadman
Professor William D. Taylor

Approved By:
Suzanne K. Damarin
Advisor
College of Education
Department of Educational Policy and Leadership
Copyright by
Sandra Kay Paxton
1996
ABSTRACT

This qualitative case study examines (1) how an administration manages the introduction of authentic assessment into a school district, and (2) the teachers' attitudes and concerns about that introduction and the implementation of authentic assessment into their classrooms, as well as the issues arising from the process. Assessment has been declared a tool for educational reform, and alternative types of assessments such as authentic assessment have been suggested as the appropriate means of testing the kinds of skills and knowledge needed for the future.

The goal of such assessments is to produce richer, clearer pictures of what knowledge the students have constructed and what they can do with that knowledge in real-world contexts. This study explores the issues, as well as benefits, encountered when introducing and implementing these new assessments. Intensive field research was conducted in a school as it initiated the transition from traditional to authentic assessment.

Findings indicate that many of the problems and issues encountered during the introduction and implementation of authentic assessment are similar to those attendant on any sort of change: resistance from teachers when ideas or actions are
initiated from central office; issues involving lack of time; communication problems between teachers, administrators, students, and parents; and issues of administrative support. Unique to changing an assessment system, or specifically to authentic assessment, are issues that have to do with reliability and validity and the paucity of models for teachers to follow.

Constructivism formed the theoretical base for this study. Many of the problems and issues encountered in this endeavor can be understood and explained from a constructivist view that everyone has different prior experiences and different interpretations of those experiences, and thus constructs his or her knowledge differently from anyone else. This poses a problem for any two or more people or groups working together, as well as for assessing learning. If everyone perceives a situation differently, how do we ever come to any consensus? The participants in this study worked through many of the problems and concerns through social negotiation; communicating and interacting with one another.
Dedicated to my mother and father,
Frances and Thomas Buckey,
and to my children, Brett and Jennifer
ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor, Suzanne K. Damarin, who supported and encouraged me throughout my program. Her insight and standards as a teacher, mentor, and friend helped me grow as a student and an individual.

I am grateful to William Taylor, a professor for whom I worked as a teaching assistant, for his confidence and friendship. He taught me to critique as well as enjoy and revere.

I wish to thank Dr. William Loadman, for his timely and valuable classes and information on authentic assessment.

I would also like to acknowledge my many friends and classmates throughout the years at Ohio State for their wealth of ideas and the camaraderie that we shared.

This dissertation would not have been possible without the generous commitment of time and energy of the teachers and administrators at the research site. I would like to extend my sincere appreciation and admiration to all who participated in this
study, especially the four teachers on the senior team and the
director of curriculum who allowed me to enter their world to
observe and ask questions, and responded graciously.

Finally, I would like to thank my family whose support,
encouragement, and love made the task a little easier. A special
thanks to my parents for always believing in me, and to Brett,
Jennifer, and Ashley who served as my inspiration during the
entire process.
VITA

April 16, 1948
Born -- Cambridge, Ohio

1973
B. S., Education
The Ohio State University
Columbus, Ohio

1973-1974
Home Instruction Teacher
Columbus Public Schools
Columbus, Ohio

1991-1995
Teaching Assistant
The Ohio State University

1993
M.A., Education
The Ohio State University
Columbus, Ohio

1994
Adjunct Professor
Ohio Dominican College

FIELDS OF STUDY

Major Field: Education
Instructional Design and Technology
Psychology
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>VITA</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTERS:</td>
<td></td>
</tr>
<tr>
<td>1. INTRODUCTION TO THE STUDY</td>
<td>1</td>
</tr>
<tr>
<td>Background of the Study</td>
<td>1</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>5</td>
</tr>
<tr>
<td>Boundaries of the Study</td>
<td>7</td>
</tr>
<tr>
<td>Definition of Terms</td>
<td>8</td>
</tr>
<tr>
<td>Significance of the Study</td>
<td>11</td>
</tr>
<tr>
<td>Summary</td>
<td>12</td>
</tr>
<tr>
<td>2. LITERATURE REVIEW</td>
<td>13</td>
</tr>
<tr>
<td>Historical Overview of Testing</td>
<td>14</td>
</tr>
<tr>
<td>Criticisms of Standardized, Objective Testing</td>
<td>17</td>
</tr>
<tr>
<td>Constructivism and</td>
<td></td>
</tr>
<tr>
<td>a Theory of Situated Learning</td>
<td>21</td>
</tr>
<tr>
<td>Alternative Assessments</td>
<td>24</td>
</tr>
<tr>
<td>Authentic Assessment Concerns and Responses</td>
<td>30</td>
</tr>
<tr>
<td>Inadequacy of Assessment Training</td>
<td>37</td>
</tr>
<tr>
<td>Research in Classroom vs Standardized Assessment</td>
<td>38</td>
</tr>
<tr>
<td>Change and Staff Development</td>
<td>40</td>
</tr>
</tbody>
</table>

viii
Staff Development and Assessment Training .......... 51
Chapter Summary ......................................................... 59

3. METHODOLOGY

Introduction ................................................................. 60
Preliminary Fieldwork ...................................................... 60
Design of the Study ......................................................... 62
Population and Setting .................................................... 64
Data Collection ............................................................... 67
Interviews ................................................................. 71
Document Analysis ......................................................... 74
Researcher's Role as Instrument ....................................... 75
Data Analysis ............................................................... 76
Data Story ................................................................. 79
Trustworthiness ............................................................ 80
Credibility ................................................................. 80
Transferability ............................................................. 83
Dependability and Confirmability ..................................... 83
Ethics .................................................................. 84

4. CONTEXT AND SEQUENCE ........................................ 86

CONTEXT ................................................................. 87
Community ................................................................. 87
School District ............................................................. 91
High School ............................................................... 96
Beginning of Senior Seminar.......................................... 103
Senior Team Teachers ................................................ 104
Ms. Dearheart ............................................................ 104
Mr. Stern ................................................................. 107
Ms. Little ................................................................. 110
Mr. Bradford ............................................................. 112
Senior Team ............................................................ 114
Grades and Ohio and District Requirements ................. 116
SEQUENCE OF EVENTS ............................................ 117
Introduction to Authentic Assessment ......................... 117
Course Postponed ....................................................... 119
Staff Meeting ........................................................... 119
Assessment Course ..................................................... 123
"Red Tape" Problems with Course ............................... 130
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff Meeting and Inservice Session</td>
<td>130</td>
</tr>
<tr>
<td>Senior Seminar</td>
<td>132</td>
</tr>
<tr>
<td>1994 Exhibition</td>
<td>135</td>
</tr>
<tr>
<td>Exhibition Grades</td>
<td>138</td>
</tr>
<tr>
<td>First Year in Retrospect</td>
<td>138</td>
</tr>
<tr>
<td>June 1994 Inservice</td>
<td>141</td>
</tr>
<tr>
<td>Second Year, 1994-1995</td>
<td>142</td>
</tr>
<tr>
<td>1995 Exhibition</td>
<td>144</td>
</tr>
</tbody>
</table>

5. FINDINGS AND DISCUSSION........................................................................148

| Question 1                                                              | 149  |
| Workshop                                                               | 152  |
| Assessment Course                                                     | 154  |
| Other Formal Introductions to Authentic Assessment                     | 162  |
| District's Future Plans for Authentic Assessment                       | 165  |
| Discussion                                                             | 168  |
| QUESTION 2                                                             | 177  |
| Teachers' Views of Assessment                                         | 178  |
| Changes for Second Year                                               | 185  |
| Administrators' Reactions to Assessment Changes                        | 195  |
| Discussion                                                             | 196  |
| QUESTION 3                                                             | 203  |
| Time                                                                   | 204  |
| Interrater Reliability and Validity                                    | 214  |
| Dealing with the Unknown                                              | 220  |
| Communication and Education of the Public                              | 222  |
| Support                                                                | 225  |
| QUESTION 4                                                             | 229  |
| Teachers' Views of Authentic Assessment                                | 229  |
| Administrators' Views of Authentic Assessment                          | 232  |
| Students' Views of Authentic Assessment                               | 233  |
| Conflicting Perceptions                                               | 239  |
| Summary                                                                | 242  |
6. CONCLUSIONS AND RECOMMENDATIONS .............................................250

   Teachers and Time .............................................................................252
   Administrative Support .....................................................................254
   Recommendations ...........................................................................255
   Changing Roles ................................................................................260
   Positive Outcomes ...........................................................................262
   Strengths of Authentic Assessment .................................................264
   Weaknesses of Authentic Assessment ..............................................264
   Understanding and Adopting a New Paradigm .................................265
   Where Are They Now? ......................................................................271
   Where Do They Go From Here? ......................................................271
   Recommendations for Further Research .......................................273

   LIST OF REFERENCES ..........................................................................275

APPENDICES

   A. ASSESSMENT COURSE SYLLABUS ..............................................294
   B. RUBRIC FOR FIRST YEAR'S PRESENTATION .............................297
   C. TEACHERS' PHILOSOPHY ............................................................299
   D. TEACHERS' EXPECTATIONS OF COURSE .................................301
   E. SECOND YEAR PROJECT RUBRIC ...............................................303
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>6 8</td>
</tr>
</tbody>
</table>

**Figure** Timeline for study...6 8
CHAPTER 1
INTRODUCTION

Background of the Study

"Technology, in its concrete, empirical meaning, refers fundamentally to systems of rationalized control over large groups of men, events, and machines by small groups of technically skilled men operating through an organized hierarchy" (McDermott, 1981, p. 142). Today, when we hear the term "technology," many of us automatically think of equipment or hardware, but according to Saettler, "the word 'technology' (the Latin form is texere, to weave or construct) does not necessarily imply the use of machines, as many seem to think, but refers to any practical art using scientific knowledge" (Saettler, 1968, pp. 5-6). Similarly, John Kenneth Galbraith defines technology as "the systematic application of scientific or other organized knowledge to practical tasks" (Galbraith, 1967, p. 12). These definitions focus on technology as a process rather than a product or machine.

Testing is a technology; it fits the definition of a tool or a process used to solve a problem. The problem may be the selection of candidates for a job or position, the selection of individuals for higher learning, the placement of individuals in a
class or program, the comparison of individuals, or the monitoring of progress of individuals or programs. The technology of testing, devised over two thousand years ago to assess behavior, has survived and continued to thrive over the millennia, amassing a colorful and interesting history. Throughout the years, the prevalent testing methodology has varied from oral formats which include interviews, presentations, debates, and discussions to written formats which include essays, short answers, true/false, and multiple-choice type items. However, both the purposes and/or forms of testing have often been subject to debate and controversy. These debates have been heard time and again, throughout the years, and so it is with the present round. Currently, the merits of "alternative" versus "objective" tests is a frequent topic in journals, conferences, faculty meetings, teachers' lounges, workshops, and even daily newspapers in many parts of the world (Broadfoot, Murphy and Torrance, 1990; Bruce, 1991; Davies, 1991; Freedman, 1995; Madaus and Kellaghan, 1993).

Testing is a critical part of the complex, integrated process called Educational Technology which involves "people, procedures, ideas, devices and organization, for analyzing problems, and devising, implementing, evaluating and managing solutions to those problems, involved in all aspects of human learning" (AECT Task Force, 1977, p. 164). In the United States, educational testing has encountered its share of debates, disputes, and reactions throughout its comparatively short, one hundred fifty
year history. For example, in response to concerns of equity and fairness, standardized tests were implemented in the 19th century (Resnick, 1982). Early in the 20th century, cost and efficiency as well as issues of fairness and subjectivity prompted studies (Starch & Elliott, 1912, 1913) which led to the promotion of more standardized, "objective," multiple-choice type testing. These "objective" tests, however, have also come under attack over the years for various reasons: critics argued that these tests discriminated against the more knowledgeable students (Hoffman, 1962), that they were biased against minorities (Medina and Neill, 1990; Trotter, 1993), that they led to the practice of "teaching to the test" thus negatively affecting systemic validity (Frederiksen, 1984; Nickerson, 1989), and that they reduced curriculum to basic skills instruction and memorization of facts and algorithms as opposed to teaching higher cognitive skills such as reasoning and problem-solving for real-life situations (Collins, 1990; Darling-Hammond, 1991; Frederiksen and Collins, 1989).

Out of the latest round of criticisms and debates, and in an effort to accommodate current interest in critical thinking and multiple perspectives, and to meet the needs and demands of colleges and employers, several new forms of testing have emerged which are quite different from the psychometric, measurement-driven tests of the last 80 years. The alternative forms, with labels such as performance-based, alternative, direct, and authentic assessment, however, are not without their share of
critics and supporters, complaints and praises, problems, and solutions, as well as cautions (Haney and Madaus, 1989; Mehrens, 1992; Wiggins, 1989a, 1989b, 1991, 1992; Worthen, 1993a). With these alternative forms of assessment, there is an attempt to make learning more meaningful, to allow students to explain and personalize their learning, to take more responsibility for their learning, and to assess their progress towards their goals. Consequently, the assessment formats used by teachers include "real-life" tasks, projects, open-ended questions, performances, exhibitions, and portfolios, in place of, or in addition to, the traditional forms of multiple-choice, true/false, and short answer items. The goal of such assessments is to produce richer, clearer pictures of what knowledge the students have constructed and what they can do with that knowledge in real-world contexts.

Research has shown that teachers who conduct the majority of assessments affecting student learning, as well as administrators who mandate, encourage, and oversee such assessments, have little preparation and need better training in classroom assessment (Conference Highlights, 1994; Schafer, 1991; Stiggins, Conklin and Bridgeford, 1986). Moreover, most studies have dealt with standardized assessment; more studies are needed which would give a clearer picture of classroom assessment (Stiggins, Conklin and Bridgeford, 1986). Also lacking in the literature are studies dealing with attitudes of students, teachers, and parents regarding classroom assessment. Although
subsequent studies by Stiggins have dealt with several of these issues, a need remains for more, in-depth studies (Stiggins and Conklin, 1992). Herman and Winters (1994) report that "the literature is silent regarding how well new practices are being implemented" (p. 55). At recent conferences (1993-1995) of the American Educational Research Association, it was evident that researchers have been attempting to fill the gap. For example, studies in Boulder, Colorado, and San Diego, California have reported findings in the areas of parent opinions, teacher practices, and the problems encountered in implementation of alternative assessments.

Purpose of the Study

The purpose of this qualitative case study is to describe the formal introduction of authentic assessment into a district, to describe the problems, issues, and concerns of teachers during the implementation of authentic assessments, and to describe the assessment practices of four high school teachers within a classroom before, during, and after their formal introduction to the ideas and practices of authentic assessment. Toward this end, I looked at classroom assessment during and after the teachers' participation in a graduate assessment course. I documented the teachers' assessment practices, their concerns and thoughts about alternative forms of assessment, and the problems, benefits, and changes resulting from the availability of, participation in, and
subsequent implementation of practices derived from the assessment course. I also examined other factors involved when using authentic assessments, as well as the expectations, attitudes and concerns of administrators throughout this process. Following the guideline of Strauss and Corbin (1990) that qualitative research questions should be open and broad, the initial questions which guided this study included:

How does a school district manage the introduction and implementation of authentic assessment into the classrooms?

What are teachers' attitudes and practices regarding assessment before, during, and after completion of an assessment course and/or other introduction to authentic assessment?

What problems or concerns and benefits arise during the introduction and implementation of authentic assessment into the district and subsequently into the classroom?

What are the strengths, weaknesses, and constraints entailed in alternative assessments, according to teachers, administrators, and students?

Informed by my early fieldwork, I had prepared the design for this study guided by the initial questions. Due to the emerging nature of qualitative research, and using the approach suggested by Bogdan and Biklen (1992), "while people conducting qualitative research may develop a focus as they collect data, they do not
approach the research with specific questions to answer or hypotheses to test" (p. 2), the original design was tentative and subject to change.

**Boundaries of the Study**

Boundaries are not always set in advance. They emerge because "boundary decisions are also an interpretive judgment based on your awareness of your data and their possibilities" (Glesne and Peshkin, 1992, p. 130). The primary foci of the study became the four teachers comprising "the senior team," their use of authentic assessment in the Senior Seminar course, and the actions, expectations, and perceptions of the administrators as they related to the introduction of authentic assessment into the district and this course. Although it is important to look at context and the factors, events, people, and policies which contribute to the attitudes, feelings, practices, behavior, and perceptions of the teachers, looking "in depth" at the analogous characteristics of administrators, parents, and students, proved too enormous a task. Although I did not study them per se, I gathered some data on students, administrators, and other teachers, as it seemed relevant and important to make sense of the four teachers' data.

The results of this research apply to the subjects, situations, and setting of this case study. The researcher was limited by the amount of time spent observing and interviewing the four
teachers, students, and administrators. Although the duration of the study was two school years, there were periods of days, weeks, and months when I had no contact with the school or teachers at all. The purpose of this study is to describe the introduction and implementation of authentic assessment with four particular teachers in the context of a specific course, so generalizations to other teachers and classrooms are impossible.

Definition of Terms

Alternative Assessment. Alternative assessment is the umbrella term which refers to the assessments under consideration here as differentiated from the traditional standardized, objective, multiple-choice tests.

Assessment. According to Gardner (1992), assessment is "the obtaining of information about the skills and potentials of individuals, with the dual goals of providing useful feedback to the individuals and useful data to the surrounding community. What distinguishes assessment from testing is the former's favoring of techniques which elicit information in the course of ordinary performance and its general uneasiness with the use of formal instruments administered in a neutral, decontextualized setting" (p. 90).
Authentic Assessment. According to Wiggins (1993b), "authentic assessment includes performance tests, observations, open-ended questions, exhibitions, interviews, and portfolios. The aim of authentic assessment is to engage students in challenges that better represent what they are likely to face as professionals and as responsible citizens. It does not measure out of context, but rather with a specific purpose in mind. For an assessment to be authentic, the context, purpose, audience, and constraints of the test should connect in some way to real world situations and problems" (p. 23). Authentic assessment applies the principle of evaluating the real work of apprentices or the challenges faced by people in the real world to students' work (Medina and Neill, 1990).

Cognitive Apprenticeship. According to Jones, Valdez, Nowakowski, and Rasmussen (1994), "While the student is engaged in challenging authentic tasks and co-investigations, apprenticeship learning takes place when the student observes, applies, and refines through practice the thinking processes used by practitioners" (p. 15).

Criteria. When using authentic assessments, it is imperative to make clear to the student what he or she is expected to do through criteria, which should be, at the least, discussed, but even better, negotiated, between the teacher and student. According to
Herman, Aschbacher, and Winters (1992), "criteria are necessary because they help you judge complex human performance in a reliable, fair, and valid manner. Scoring criteria guide your judgments and make public to students, parents, and others the basis for these judgments" (p. 45).

**Exit Exhibition.** The exit exhibition is the culminating experience of the seniors in the study. At the public exhibition, the seniors, whose presentations were judged by their teachers to be the best of the class, present to a broader audience of peers, parents, teachers, and community members as well as a panel of judges.

**Fee Waiver.** In an arrangement between a school system and a college, teachers are allowed to attend classes at the college using fee waivers instead of actual money, in exchange for, the use of the classrooms within the district by student teachers and researchers from the college.

**Rubric.** A rubric is a type of scale used to rate the presence or quality of a performance or product. According to Herman, Aschbacher, and Winters (1992), "there are three major types of scales: checklists, numerical ratings, and qualitative (either descriptive or evaluative) ratings" (p. 64).
Senior Seminar. The Senior Seminar is a course required for graduation from the district in this study. It was designed to help the seniors in their transition from high school to the "real world" through four phases: team building, community service, portfolios, and project and presentation.

Senior Team. The senior team is made up of four high school teachers from different disciplines: English, social studies, mathematics, and science, who team teach the Senior Seminar course.

Significance of the Study

As education moves in the direction of preparing students for a world where information management, critical thinking, and problem solving skills are crucial, finding a means to assess such skills appropriately becomes a major challenge. In this study, the researcher looked at the introduction and implementation of authentic assessment which many claim can meet that challenge. Moreover, as authentic assessment necessitates the inclusion of students and teachers in the preparation, implementation, and grading process, the importance of training teachers in assessment theory and procedures also becomes apparent. Therefore, the study also looked at attempts by school administrators to provide such staff development in authentic assessment in order to discover problems, concerns, and/or benefits which are
encountered from the perspectives of the teachers and administrators. This may provide suggestions as well as cautions to other districts proposing similar programs.

**Summary**

In the first chapter, I have surveyed some of the problems and debates related to objective testing, and have described the alternative forms of assessment that have emerged in an effort to meet current needs. Other concerns have been outlined which include: lack of teacher preparation in assessment matters and lack of research in the area of classroom assessment practices as well as attitudes and issues concerning such assessments. In addition, the purpose of the study is stated and the research questions which guided it are listed, the boundaries of the study are set forth, terms relevant to the understanding of the study are defined, and the significance of the study presented.

Chapter 2 examines how other research has addressed issues related to this study. The methodological design of the study is outlined in Chapter 3 while the context and sequence of events relevant to this study are described in Chapter 4. In Chapter 5, the research questions are addressed and the data analyzed and reported. Finally, in Chapter 6, the conclusions and recommendations for further studies are presented.
CHAPTER 2
LITERATURE REVIEW

Politicians and educators have indicated that educational reform is necessary to deal with the changing workplace, the technological revolution, and the information explosion in this country. Several considerations are required: the kinds of skills and knowledge that will be needed in the 21st century, how workers will deal with the changing challenges, who will be responsible for student learning, and many others. Assessment has been mentioned as a potentially effective tool for such reform, and alternative types of assessments have been suggested as the appropriate means of testing the kinds of skills and knowledge needed for the future. This case study examines (1) how an administration manages the introduction of authentic assessment into a school district, and (2) the teachers' attitudes and concerns about that introduction and the implementation of authentic assessment into their classrooms, as well as the issues and problems arising from the process.

Several areas of research are pertinent to this study. An historical overview of testing is useful in order to understand how and why schools and teachers currently use testing as they do. It
is necessary to look at what research in psychology and sociology has shown about learning, especially the kinds of learning and knowledge that are essential for problem solving and reasoning, for locating, selecting, and using information, and life-long learning and self-assessing. This is followed by an examination of alternatives which have been suggested for assessing the kind of learning necessary in the future with a consideration of their strengths and weaknesses. Teacher education in testing is another area which needs to be reviewed to find out what teachers do and do not learn about assessment in their education preparation courses. Looking at ways in which new ideas and programs are introduced to teachers through staff development after they enter the teaching profession is considered. Finally, studies that have dealt specifically with staff development in authentic assessment are discussed.

**Historical Overview of Testing**

Standardized, "objective" tests have proliferated since the early part of this century as a means of sorting and ranking the growing numbers of students in public schools. In 1900, the eighth grade promotion test was abolished in U.S. schools allowing more students to stay in school longer. In addition, large numbers of immigrants were entering the United States. As a result of these factors, enrollment in elementary and secondary schools increased by 20% overall between 1910 and 1920, while the
secondary school enrollment alone increased by 140% (Haney, 1984). The public schools were suddenly expected to teach and sort a vast number of students from extremely diverse backgrounds. Ideas outlined by Frederick Taylor in *The Principles of Scientific Management* (1911) provided a model of efficiency and economy that schools could use as a means to accommodate the burgeoning numbers of students.

Until the early part of this century, tests were predominantly essay exams, and were tedious and time consuming to grade. Studies conducted in 1912 and 1913 by Starch and Elliott brought into question the reliability of essay tests on the basis of the discrepancies they found among teacher judgments and the variation in the standards used for scoring. It was around this same time that a more efficient means in the form of "objective," paper and pencil tests was developed to test the large numbers of World War I recruits. This major breakthrough was seen as the solution to the problems of both subjectivity and efficiency in public school testing.

According to Stiggins (1991b, 1993), a testing community developed which for the next 60 years took over from classroom teachers the role of test preparer, grader, and interpreter. Testing became highly sophisticated and complex, and much too complicated for the average teacher. Consequently, classroom teachers and administrators gladly "abdicated responsibility for understanding or conducting assessments in schools, leaving it to
test and textbook publishers to develop 'scientifically precise' assessment tools" (Stiggins 1991b, p. 265). This abdication, however, also led to the separation of instruction and assessment (Stiggins, 1991b).

The initial purpose of these tests was to differentiate and sort students efficiently and effectively for their ultimate roles, that is, factory worker, college student, and so forth (Darling-Hammond, 1991; Stiggins, 1993). Over the next decades, the use of these tests changed from sorting and selecting on a broader student, teacher, and system level, to accountability, first on the local level, then the state level, and finally on the national level (Stiggins, 1993). Testing for accountability is one of the major reasons that Resnick and Resnick (1985) refer to American school children as "the most tested in the world and the least examined" (p. 17). According to Resnick and Resnick (1992):

At present, accountability testing has a more profound influence on educational practice and possibilities than testing for either selection or instructional management. That influence derives from the growing pervasiveness of accountability testing in American education and the control - direct and indirect - that mandated accountability tests exert over curriculum and teaching practice at all levels of the school system. (p. 52)

Gardner (1992) sums up the situation with his statement, "The United States is well on the way to becoming a 'complete testing society'" (p. 79). 

16
Criticisms of Standardized, Objective Testing

Since the 1970s, a renewed movement for accountability has reinforced and promoted the use of standardized, objective paper and pencil tests. Archbald and Newmann (1988) and Haney and Madaus (1989) point out that the accelerated use of these "objective" remedies has produced many criticisms and complaints. For example, one consequence of the use of these tests for accountability has been the widespread practice of "teaching to the test" which effects systemic validity (Frederiksen, 1984; Nickerson, 1989; Resnick and Resnick, 1985). Furthermore, teaching to the test reduces curriculum to basic skills instruction and memorization of facts and algorithms as opposed to the teaching of higher cognitive skills such as reasoning and problem solving for real-life situations (Collins, 1990; Darling-Hammond, 1991; Herman, 1992; Herman, Aschbacher, and Winters, 1992; Lieberman, 1991; Shepard, 1989).

In the early part of the century, when the public schools were given the task of educating the masses to prepare them for jobs which required little thinking, this type of instruction and testing may have been effective. At that time, development of higher order thinking was reserved for the elite who attended private or religious schools (L. Resnick, 1987) and for post secondary education. However, as we move into an era when problem solving and critical thinking skills are essential for everyone, and when it is not enough to memorize and recite facts,
a different type of instruction and testing is necessary (Herman, Aschbacher and Winters, 1992; Resnick and Resnick, 1992).

In the past, instruction was based on a behaviorist or objectivist model of learning in which it was thought that information or knowledge could be broken down into smaller pieces and transferred from teacher to student. Standardized, objective tests provided a means to assess that type of linear, rote learning. However, many educators have come to subscribe to a theory of learning which is based on studies in cognitive psychology, cognitive science, and sociology. Research has shown that individuals construct their own knowledge based on past experiences and their interpretation of those experiences, and that cultural and social factors as well as context and activity play important roles in learning (Brown, Collins and Duguid, 1989; Lave, 1988; Resnick, 1989; Schoenfeld, 1985; Suchman, 1990). Based on this theory of learning, a different kind of instruction, and subsequently, a different type of testing are needed.

Standardized, objective tests were not created for and are not suitable for diagnosis and remediation, but rather for accountability (Resnick and Resnick, 1985). Psychometric theory may work to differentiate individuals and may even provide some predictive power, but it is not appropriate for diagnostic testing, that is, testing to determine strengths, weaknesses, and misunderstandings or misconceptions (Linn, 1986; Siegler, 1989; Wiggins, 1993a; Wittrock, 1991b). Standardized tests merely
point out that a particular piece of information is not known or understood, but offer no clarification concerning what is missing, how it is misconstrued, or what should be done about it. Wiggins (1993a) emphasizes that "successful learning depends upon adjustment in response to feedback; no task worth mastering can be done right on the first try" (p. 12). Mislevy (1993) points out that the psychometric theory on which most current tests are based has not kept up with research and findings in relevant areas; he states, "It is only a slight exaggeration to describe the test theory that dominates educational measurement today as the application of 20th century statistics to 19th century psychology" (p. 19). He elaborates, "The essential problem is that the view of human abilities implicit in standard test theory ... is incompatible with the view rapidly emerging from cognitive and educational psychology" (p. 19). Linn (1986) shares this view and suggests that research in cognitive psychology may provide direction in this area, and that "testing needs to be more closely linked to instruction" (p. 1159). Likewise, Wiggins (1993a) contends that "the stubborn problems in assessment reform have to do with a pervasive thoughtlessness about testing and a failure to understand the relationship between assessment and learning" (p. 3).

In addition, researchers have reported that students do not transfer in-school learning to real-life situations (Brown, Collins and Duguid, 1989; Lave, 1988; L. Resnick, 1987; Wittrock, 1991a).
Research suggests that in-school learning and assessment must resemble situations and problems which students will encounter in the real world. Collin's explains (1990):

Otherwise students are not likely to see how the knowledge they are getting can be applied. The 'what' of knowledge is only a third of what needs to be learned; we also need to know the 'when' and 'how' it applies in different contexts, else we will find that students cannot transfer what they have learned to new, but relevant, contexts. (p. 77)

We need problems or tasks which are "ill-structured" (Frederiksen, 1984) or "messy" (Wiggins, 1993a) rather than contrived, decontextualized test items.

In addition to their effects on programs and the kinds of things that are taught and learned, traditional tests also have a direct impact on how the student sees him/herself. Collins (1990) cites a study by Dweck (1986) which found that students who did not perform well on tests developed negative self-images, lost their motivation to learn, and came to view school as irrelevant and boring.

Besides affecting instruction, these objective, paper and pencil tests have adverse effects on learning as indicated by Bennett (1993). He cites studies which found "students to prefer taking multiple-choice measures even though they regarded the constructed response as a more accurate indicator" (p. 18).

Other criticisms are that standardized, objective tests give inaccurate information about what learning occurs in the nation's
schools (Medina and Neill, 1990), and that they are biased or unfair to those other than white, middle-class males (Medina and Neill, 1990; Rosner, 1993; Trotter, 1993; Wolf, 1993b).

Constructivism and a Theory of Situated Learning
Studies by Bednar, Cunningham, Duffy and Perry (1992), Brown, Collins and Duguid (1989), Duffy and Jonassen (1992), Lave (1988), and Suchman (1990) have contributed to a challenge to the objectivist theory of learning and the development of an alternative explanation called a theory of situated learning and cognition and a view of knowledge known as constructivism.

Constructivists believe that we all create our own knowledge based upon experiences and our interpretation of those experiences. Meaning is rooted in and grows through experiences. The internal representation of knowledge of each person is constantly being added to and changed. Constructivists emphasize the importance of context and meaning, and agree, in general, with a theory of situated learning (Bednar, Cunningham, Duffy, and Perry, 1992).

According to Brown, Collins, and Duguid (1989), Lave (1988), and Suchman (1990), situated learning is a paradigm of learning which takes into consideration "where the learner is" and what prior experiences and knowledge he/she brings to the learning situation. It assumes that "knowledge is situated, being in part a product of the activity, context, and culture in which it is
developed and used" (Brown, Collins, and Duguid, 1989, p. 32). It assumes that learners do not follow plans in order to learn, but rather learn by means of face-to-face communication, multiple contextualized experiences, collaborative interactions, and the sharing of multiple perspectives (Streibel, 1989).

It is through sharing multiple perspectives and collaborative interactions that students are able to make sense of the new information and assimilate or accommodate it into their existing internal representation of knowledge. It is through this same communication and interaction and looking at how students process information that teachers sense misconceptions or misunderstandings of students and are able to improve instruction, and coach and guide students (Wiggins, 1993a; Wittrock, 1991a).

Practices based upon situated learning provide for use of multiple contexts, including in-school and out-of-school problem situations and collaborations of real-life or authentic activities and practice (Pea, 1988, p. 21). Thus, knowledge is presented within the context in which it is used, and the student becomes an active participant in his/her learning. Such practices utilize a "learn by doing" format which allows the student to enter a "cognitive apprenticeship" which supports construction of personal knowledge through interactions with the teacher or others. This makes the learning meaningful and establishes purpose to the learning.
The teacher's role in situated learning becomes that of mentor. The teacher must be willing to give up the role of information provider and manager, and accept the role of guide, facilitator, and coach. The goal of the teacher should be to involve the students in their own learning and wean the students to become self-teachers, self-assessors, and life-long learners.

The teacher should also be concerned with creating a learning environment free from "arrogant perception." According to Damarin (1993), in this context, "perception' refers to the knowledge/attitude with which a person meets a situation. A person's perception (of another or of a situation) is arrogant when it exists or is expressed in subtle or overt ways which devalue the other or the situation" (p. 28). When arrogant perception is present, collaboration, exploration, and cooperation necessary for building situated knowledge is thwarted. Therefore, it should be the responsibility of the teacher to provide a safe space for knowledge construction to occur, free from sexism, racism, and classism, and also free from the "world views such as scientism, institutional practices such as academic tracking, and production systems such as instructional design [which] also lend themselves to arrogant perception" (Damarin, 1993, p. 28). Damarin (1994) suggests that "the implementation of theories of situated cognition and learning has the potential to make education more open to students from diverse backgrounds" (p. 19) and to empower all students.
It should be apparent that instruction and practices based upon a theory of situated learning and constructivism will be much different from traditional instruction which operates under the assumption that the teacher possesses the knowledge and transfers it to the student. Consequently, relying solely on objective tests would not be an appropriate means of assessing student learning within this model.

**Alternative Assessments**

Assessment has been proposed as one of the means for educational reform in the 1990s (Herman, Aschbacher and Winters, 1992; Simmons and Resnick, 1993). According to Resnick and Resnick (1992), "Tied to curriculum and designed to be taught-to, performance assessments can become essential tools in educational reform" (p. 72). Wolf, LeMahieu, and Eresh (1992) have suggested that setting high standards and using different kinds of assessment can serve as tools for educational reform. Educators and psychologists have posited some suggestions of assessments that promote the kind of learning which seems more appropriate for this new informational age in which knowledge is exploding at a staggering rate and which requires reasoning and problem solving. Since problem solving cannot easily be assessed by multiple-choice, decontextualized questions, but rather, needs to be assessed by meaningful interactions which require thinking and analyzing in specific situations, Frederiksen (1984) proposes
giving students ill-structured problems to solve; the kind that are found in real life. Frederiksen and Collins (1989) suggest "developing tests that directly reflect and support the development of the aptitudes and traits they are supposed to measure" (p. 28). They envision a type of assessment that would require a number of tasks to be performed, some of which would assess students' understanding of critical concepts and would be open ended, thus permitting students to express special knowledge and creativity (p. 31). Archbald and Newmann (1988) discuss authentic academic achievement and how it should "integrate knowledge in two ways. Not only must students be challenged to understand integrated forms of knowledge, they must also be involved in the production, not simply the reproduction, of new knowledge, because this requires knowledge integration" (p. 3). Wiggins, (1989a, 1991, 1992) similarly, recommends designing assessment tasks which are authentic, meaningful, enabling, challenging, messy or ill-structured, contextually situated, have set standards, and often require collaboration. He maintains that the first step in designing tests should be to decide "what are the actual performances that we want the students to be good at" (Wiggins 1989a, p. 705). Medina and Neill (1990) agree that "authentic evaluations indicate what we value by directing instruction toward what we want the student to know and be able to do" (p. 58). According to Rudner and Boston (1994), "The performance assessment movement
encompasses much more than a technology for testing students. It requires examining the purposes of education, identifying skills we want students to master, and empowering students" (p. 2).

Alternatives to standardized, objective testing under the labels of "authentic," "direct," "performance," and "alternative" assessment have been proposed and are being implemented, either for certain subject areas such as writing or for the whole curricula. Alternatives are being piloted or used in many individual schools and districts; by entire states (for example, Vermont, Connecticut, California, Kentucky); and by professional associations for subject areas such as National Council of Teachers of Mathematics. Also trying alternative assessments, are schools involved in research projects with programs such as the Coalition of Essential Schools, the Center for Research on Evaluation, Standards, and Student Testing (CRESST), the California Assessment Collaborative (Darling-Hammond, 1991; Martin-Kniep and Kniep, 1992; Mitchell, 1992).

The terms or labels to describe these assessments are often used interchangeably, but it has been recommended (Meyer, 1993; Wiggins, 1993b; Worthen, 1993a) that distinctions be made clear in an attempt to educate the public as to the respective characteristics of each. Wolf (1993a) distinguishes between testing and assessment:

**Standardized testing** was designed to be a post-hoc episode of measurement, independent of the curriculum, serving chiefly the purposes of certification and
accountability. **Assessment** is meant to be an episode in which both students and teachers learn about the standards of good work and how to achieve them. (p. 237)

Gardner (1992), also, voices a concern with testing, while portraying a more positive view of assessment:

> I define assessment as the obtaining of information about the skills and potentials of individuals, with the dual goals of providing useful feedback to the individuals and useful data to the surrounding community. What distinguishes assessment from testing is the former's favoring of techniques which elicit information in the course of ordinary performance and its general uneasiness with the use of formal instruments administered in a neutral, decontextualized setting. (p. 90)

**Alternative assessment** is the umbrella term which differentiates assessments from the traditional standardized, objective, multiple-choice tests. According to Wiggins (1993b):

**Performance-based** assessment allows teachers to evaluate a student's skill by asking the student to perform tasks that require the skill. The student is to perform with knowledge instead of merely recalling or recognizing other people's knowledge. **Authentic** assessment includes performance tests, observations, open-ended questions, exhibitions, interviews, and portfolios. The aim of authentic assessment is to engage students in challenges that better represent what they are likely to face as professionals and as responsible citizens. It does not measure out of context, but rather with a specific purpose in mind. For an assessment to be authentic, the context, purpose, audience, and constraints of the test should connect in some way to real world situations and problems. (p. 23)
"Authentic assessment" and the associated idea of "cognitive apprenticeship" are derived from the examination of the work of craft apprentices (L. Resnick, 1989). In the process of becoming a glass maker or an artist, three basic elements are crucial: observation, coaching, and practice (Collins, Brown, and Newman, 1989). According to Jones, Valdez, Nowakowski, and Rasmussen (1994):

While the student is engaged in challenging authentic tasks and co-investigations, apprenticeship learning takes place when the student observes, applies, and refines through practice the thinking processes used by practitioners. Students can then reflect on their practice in diverse situations and across a range of tasks and articulate the common elements of their experiences. This will enable them to generalize their skills and transfer their learning to new situations. (p. 15)

Authentic assessment applies the principle of evaluating the real work of apprentices or the challenges faced by people in the real world to students' work (Medina and Neill, 1990). In the classroom, authentic assessment encompasses a variety of techniques rather than relying solely on the numerical results of standardized, objective tests.

Authentic assessment borrows from the craft apprentice's work in another way. The apprentice is provided with a concrete model to insure high quality craftsmanship. Likewise, the student is provided with exemplars of good and bad work which act as models and motivation. Thus, authentic assessment sets specific guidelines and standards which students must internalize and
strive toward. Worthen (1993a) notes that "insightful teachers recognize that alternative assessment has long constituted the core of their methods for assessing student learning in the classroom" (p. 445), and that "our nation's classrooms have been quietly awash in such performance-based assessment for decades" (p. 445).

The issue of teaching to the test has been blamed for interfering with the predictive validity of tests, but Wiggins (1989b) argues that with authentic assessment, teaching to the test does not jeopardize validity. The assessments work like this: teachers model the desired skills that students then claim as their own and practice. The criteria for performance on these tests are specifically defined, by the teacher or other authority and sometimes through collaboration between the teacher and students, but, at the least, they must be discussed with and made clear to the students. "Criteria are necessary because they help you judge complex human performance in a reliable, fair, and valid manner. Scoring criteria guide your judgments and make public to students, parents, and others the basis for these judgments" (Herman, Aschbacher, and Winters, 1992, p. 45). Exemplars of expectations are also available for the students to review and consider. Teachers acting as coaches help the students to understand the criteria and to practice skills, self-assessment, and reflection. The testing becomes an integral part of instruction, that is, a means to the end of actual performance of the new skills
whether problem solving, writing, or speaking a new language, and the "test" measures what it is intended to measure, thus insuring validity.

The relationship between student, teacher, learning, and instruction is very important. Assessment should not exist outside this relationship (R. Brown, 1989). A problem with this is that the 60 year era of psychometrics, which separated the roles of teaching and testing, left teachers unprepared to pick up the testing role. Although they conduct daily assessments in their classrooms, most lack knowledge concerning reliability, validity and interpreting scores (Stiggins, 1993).

Authentic Assessment Concerns and Responses

Authentic assessment is not without its critics and/or skeptics, however. Mehrens (1992) defends the use of multiple-choice testing and argues that performance assessments used for accountability are time intensive, expensive, and have validity and reliability problems. First of all, Mehrens states that the criticism that multiple-choice items cannot test higher order thinking skills is invalid. Others agree that it is possible to write items that get at higher level thinking, but that few actually do (Frederikson, 1984; Herman, 1991). Herman (1991) makes a strong case that knowing "how" and "when" information should be used, and in what contexts, is critical to learning and is difficult to assess with decontextualized multiple-choice items.
Mehrens (1992) argues that performance assessments take considerable time to create and once they have been used they may not be used again. Baron (1991) suggests creating new ones or borrowing from other teachers, schools, or districts. She points out that if each teacher created only one performance task each year which could be shared with other teachers, our only problem might be developing the means to exchange all the possibilities.

Concerning reliability, Mehrens (1992) claims that there are several problems. He contends that the threats to reliability have to do with "the small number of independent observations (the sampling problem...). A second has to do with the subjectivity of the scoring process" (p. 7). Wiggins (1993a) addresses subjectivity by quoting Benjamin Bloom:

"Synthesis is a type of divergent thinking: it is unlikely that the right solution can be fixed in advance." Higher order assessment will therefore almost always be judgment-based. "Each student may provide a unique response to the questions or problems posed. It is the task of the evaluator to determine the merits of the responses" (We must develop apt scoring criteria and standards that reward diverse excellence, in other words). (pp. 10-11)

A third problem, according to Mehrens, has to do with a lack of internal consistency that influences generalizability" (p. 7).

Of these three problems, interrater reliability has received the most attention. It is argued that if the grading criteria, scales, and or rubrics are explicit and the graders are well trained, the interrater reliability can be increased to an acceptable level.
(Frederikson and Collins, 1989; Herman and Winters, 1994; Linn and Burton, 1994). This has been shown with portfolios and writing samples which have been used for several years in many schools and have proven to be successful in improving student work and attitudes toward learning (Mitchell, 1992; Suhor, 1985; Tierney, Carter, and Desai, 1991). Research is continuing on the development of performance criteria and how to achieve better interrater reliability through training (Baker, Freeman, and Clayton, 1991; Perlman, 1993).

The most well known study of interrater reliability has been that of the Vermont portfolio program. In 1988, Vermont implemented a statewide performance assessment program "to provide high-quality data about student achievement (in this case, sufficient to permit comparisons of schools or districts) and to induce improvement of instruction" (Koretz, Stecher, Klein and McCaffrey, 1994, p. 5). Reports for 1991-92 revealed very low interrater reliability for both writing and math portfolios, and had psychometricians and public officials shaking their heads and questioning the usefulness of such assessments. The rater reliability for math improved significantly in 1993, but not for writing. However, Koretz and colleagues who compiled and analyzed the data, caution that the low rater reliability achieved in the portfolio assessments may be misleading due to the unique aspects of the program. They believe that one factor effecting the reliability was the lack of standardization of tasks and the
problem with training a large number of raters to a sufficient level of accuracy. As a result, the program has been unsuccessful in meeting its first purpose of "providing high quality data about student performance" (p. 11). The program may have been more successful in achieving its second purpose of improving instruction because it served as a strong tool for changing instructional practices, however, there was incomplete "understanding of the desired types of practice" (p. 13). In addition, the cost was great in terms of money, time, and stress. Koretz and colleagues maintain that "the basic lesson to be drawn from its experience is the need for modest expectations, patience, and ongoing evaluation in our national experimentation with innovative large-scale performance assessments as a tool of educational reform" (p. 15).

Herman and Winters (1994) report that interrater reliability is achievable but that the bigger problem remains in the costs of achieving it, that is "staff training, development of task specifications and prompts, administration of portfolio records and their storage and scoring" (p. 54), as well as "opportunities for professional development, ongoing support, technical expertise and time for teachers to develop, practice, reflect upon, and hone their instructional and assessment expertise" (p. 54). They further report that "thus far, the literature is silent regarding how well new practices are being implemented" (p. 55).
Several studies have shown that high interrater reliability scores can be achieved (Baker, Freeman, and Clayton, 1991; Herman and Winters, 1994). However, Linn and Burton (1994) contend that the bigger problem "is the relatively limited degree of across task generalizability" (p. 5). They suggest the need for a greater number of tasks to reach a sufficient confidence level for generalizability, and also "the need to use information obtained about student performance throughout the year rather than relying on a single examination" (p. 6). They claim that this is extremely important when the results of the assessments are used "to make pass-fail certification decisions for individual students" (p. 5). Brennan and Johnson (1995) have also addressed this problem and warn that "assessments requiring considerable time may result in student scores that are not dependable if each score is based on few tasks and if student-task variance is large" (p. 12).

The validity issue was mentioned earlier in the discussion of teaching to the test. It has also been discussed by Frederiksen and Collins (1989) who expanded "the notion of construct validity of a test to take into account the effects of instructional changes brought about by the introduction of the test into an educational system" (p. 27). They refer to this concept as systemic validity and argue that tests should be developed which measure what it is we value and what we want students to know and "thus reflect and support the development of the aptitudes and traits they are supposed to measure" (p.28). The "direct tests" they suggest have
been described earlier, and reflect the same characteristics as authentic assessments. They have proposed a set of design principles for creating tests which are systemically valid. The standards for these must include directness, scope, reliability, and transparency.

Linn, Baker, and Dunbar (1991) have also addressed the issue of judging the quality of assessments, and suggest that traditional criteria need to be expanded. They claim that "reliability has too often been overemphasized at the expense of validity; validity itself has been viewed too narrowly" (p. 16). They propose eight additional criteria: consequences, fairness, cognitive complexity, content quality, content coverage, meaningfulness, and justification of cost. Discussions of these criteria and their importance in evaluating authentic assessment procedures have recently appeared in journals for teachers and administrators, (for example, Bracey, 1993) in an effort to alert them to the kinds of criteria to consider when preparing authentic assessments. In a recent article by Baker (1994), she asserts that:

Although opponents to new assessments raise credible questions of cost, feasibility, objectivity, and fairness, for the most part educators are embracing what they believe to be a more useful view of assessment. The support of measurement experts - who have shifted their attention from reliability coefficients to the consequences of measures used to evaluate student achievement - further bolsters arguments in favor of performance assessment. (p. 58)
There is still serious debate about whether authentic assessment can be used solely as a diagnostic and instructional tool or whether it can also serve an accountability role. Whatever the outcome of this debate, the advent of authentic assessment will necessitate several changes in the functions of teachers, students, and testing organizations. The role of testing companies will be quite different as the "practice of assessment will no longer be confined to the testing organizations; it will become more decentralized, as teachers and students are taught to internalize the standards of performance for which they strive" (Frederiksen and Collins, 1989, p. 32). According to Cizek (1993), in this new model, "testing specialists function as consultants and advisors to curriculum experts and teachers ... the role is clearly supportive, not directive" (p. 8). Shepard (1993) adds, "Content experts and psychometricians should work collaboratively" (p.12), while Worthen (1993a) cautions:

There is an element of risk if large testing companies are allowed to become the primary source of instruments for alternative assessment. Sole dependence on measures developed in isolation from local curricula would clearly undermine one of alternative assessment's greatest strengths: the integration of instruction and assessment. (p. 453)

However, additional time, support, training, and resources will need to be allocated to teachers conducting such assessments (Herman and Winters, 1994). Thus, in order to gain these needed resources and support, it will also be important to expand the
concept of testing through Stiggin's idea of assessment literacy for parents, administrators, and the general public while informing and educating them of the positive value and long range aspects of authentic assessment.

**Inadequacy of Assessment Training**

Research on programs and agencies responsible for certifying teachers shows that the majority of teachers have not received adequate, if any, assessment training. "The vast majority of teachers entering the profession are deficient in measurement training" (O'Sullivan and Chalnick 1991, p. 18). According to Schafer (1991) citing Schafer and Lissitz (1987), "only about half of the teacher education programs in the nation require a course in measurement for initial certification" (p. 3) and, later citing Gullickson (1986), "when such a course is included in a student's program, it often does not include what is important for teachers to know as judged by teachers themselves" (p. 3). Stiggins (1991a) reveals the breadth of the inadequacy by discussing our "assessment literacy" as a nation. He asks: "Who prepares students to understand feedback? Who trains parents to interpret and understand test scores? What training is received by school boards, legislators and other public officials who set policies regarding testing?" (p. 3). He cites the dismal figures reported by Schafer & Lissitz and Gullickson which reveal that so few teachers receive courses in measurement, but goes beyond them to assert
that administrators are even less well trained in assessment than teachers (Stiggins, 1991a; Stiggins, 1995). These reports that teachers lack adequate assessment training combined with the fact that 99% of all assessment events in American schools today are conducted by classroom teachers, while millions of dollars are spent on the remaining 1% large-scale assessments (Stiggins 1991b; Stiggins and Conklin, 1992), point out a serious discrepancy, and the need for reevaluation of priorities. As Stiggins reminds us, it is the classroom assessments "that provide the energy that fuels classroom practice" (1994, p. viii). Plake and Impara (1993) report findings of their research on teacher assessment literacy. They administered an instrument to assess areas of teacher competency in measurement to 555 teachers and 286 administrators. The teachers scored highest "on the subscale measuring teacher knowledge in the area of administering, scoring and interpreting test results; poorest performance was on items measuring the teachers' knowledge about communicating test results" (p. 12). Asked about their preference for learning to interpret test scores, "nearly 60% chose an in-service delivery option" (p. 12).

**Research in Classroom vs Standardized Assessment**

Research further shows that most studies having to do with assessment have dealt with standardized testing; few have dealt with classroom assessment which most directly effects students
and learning. Stiggins and Conklin (1992) maintain that "it is teachers' classroom assessments that students rely on to help them set expectations of themselves, learn what teachers expect, practice hitting valued targets and decide whether to care" (p. 208-209). In research of teacher attitudes concerning testing, Stiggins, Conklin, and Bridgeford (1986) found that "by far the most frequently expressed concern was uncertainty about how to improve test quality and manage the assessment environment" (p. 9). They point out that "teachers trained only in paper and pencil measurement methods face real difficulties in the classroom assessment environment" (Stiggins et al., 1986, p. 14) where oral questioning, performances, checklists, interviews, and observations play such important roles in assessment.

The implications of their studies indicate the need for (1) "in depth studies of classroom assessment environments" (Stiggins et al., 1986, p. 15), and (2) training which includes "measuring higher order reasoning skills, writing quality paper and pencil test items, integrating assessment and instruction through oral questioning strategies, and designing quality performance assessments based on observation and professional judgment" (Stiggins et al., 1986, p. 15). Moreover, Herman, Aschbacher, and Winters (1992) lament that "despite numerous conferences and meetings on these topics [need for reform and emphasis on assessment], educators have had little concrete guidance in the creation and use of alternative assessments" (p. 1).
Change and Staff Development

Commensurate with the need to change the kinds of assessments that have been used in schools is the need to provide the means for teachers and administrators to learn about the new approaches and how to implement alternative assessments. Change can be perceived as an exciting challenge or an ominous threat to the staff of a school. Deal and Peterson (1990) claim that in order for reform or change to be effective, administrators "must work simultaneously on staff needs and skills, the organization's goals and roles, and the dynamics of political power and conflict" (p. 7). Beyond that, they address another element crucial to the success of any reform, that is, the acknowledgment of and willingness to work within the "school culture" which they describe as "the character of a school as it reflects deep patterns of values, beliefs, and traditions that have been formed over the course of its history" (p. 7). They caution that neglect of the symbols and symbolic activity related to this "culture" can result in "reductions in cooperation, productivity, commitment, and motivation" (p. 9).

If schools are to incorporate new ways of teaching and assessing that promote student learning, administrators and teachers must have the opportunity to learn about these new ideas. Fullan (with Stiegelbauer, 1991) defines professional development as "the sum total of formal and informal learning experiences throughout one's career from preservice teacher
education to retirement" (p. 326). Since my study deals with educators who are already out in the schools, the term staff development will be used to refer to this aspect of professional development.

Sparks (1994) discusses the paradigm shift that must occur in staff development. He identifies the three ideas which are altering education today as: results-driven education, systems thinking, and constructivism. Sparks suggests that staff development must adopt these same changes while the goal should be improved performances, not time spent in workshops. He also stresses that teachers should be allowed to construct the solutions to problems through collaboration with peers, researchers and experts via discussions, activities and reflection on the job rather than being given the answers through seminars.

Theorists of educational change agree that teachers must see the need for change, and must be committed to and feel ownership of new programs and practices in order for them to succeed (Dillion-Peterson, 1990; Fullan with Stiegelbauer, 1991; Goldenberg and Gallimore, 1991; Shanker, 1990; Wood and Thompson, 1993). A survey by McBride, Reed, and Dollar (1994) concluded that the attitude with which a teacher approaches staff development ultimately effects what and how much change is brought about or if the new ideas and concepts will even be adopted at all. One finding indicated that 70% of the teachers "believed that commitment to staff development would be greater
if more teachers were involved in the planning process" (p. 40). Similarly, Dillon-Peterson (1990) report the importance of involving representatives of the target audience in the needs assessment, planning, execution, and follow-up of any staff development activities; "even in instances where mandates are necessary, great care should be taken to respect the wants and needs of individual staff members" (p. 224). According to Shanker (1990), "Teachers are viewed as an important source of knowledge that should inform what happens in schools" (p. 93) while Wallace, LeMahieu and Bickel (1990) emphasize that "the program must be responsive to the needs of those it serves and consonant with a vision of educational reform and change" (p. 188). Fullan (with Stiegelbauer, 1991) cautions that "the extent to which proposals for change are defined according to only one person's or one group's reality (e.g., the policy-maker's or administrator's) is the extent to which they will encounter problems in implementation" (p. 36).

Workshops and inservices have traditionally been the primary means for staff development. However, utilizing theories of learning based upon the idea that learners construct their own knowledge in a particular context and according to their past experiences and their needs, a different view of staff development has emerged. According to Lieberman (1995), "processes, practices, and policies that are built on this view of learning are at the heart of a more expanded view of teacher development that
encourages teachers to involve themselves as learners - in much the same way as they wish their students would" (p. 592).
Darling-Hammond and McLaughlin (1995) concur with this view and suggest that "professional development today also means providing occasions for teachers to reflect critically on their practice and to fashion new knowledge and beliefs about content, pedagogy, and learning" (p. 597). They see the role of administration as being supportive by "rethinking schedules, staffing patterns, and grouping arrangements to create blocks of time for teachers to work and learn together" (p. 601). In addition, they also suggest grouping teachers into teams so that planning for the students' learning experiences can be coordinated and shared by all the teachers involved.

When implementing change, Fullan (with Stiegelbauer, 1991) suggests three aspects of an innovation which he contends are necessary for its success:

(1) the possible use of new or revised materials (direct instructional resources such as curriculum materials or technologies), (2) the possible use of new teaching approaches (i.e., new teaching strategies of activities), and (3) the possible alteration of beliefs (e.g., pedagogical assumptions and theories underlying particular new policies or programs). (p. 37)

Further recommendations have been made as to how staff development can serve the purpose of introducing and implementing change. Goldenberg and Gallimore (1991) suggest that "staff development ... must be grounded in the mundane but
very real details of teachers' daily work lives and in a form that provides the intellectual stimulation of a graduate seminar" (p. 69-70). The requirements they see as necessary for change to occur are: teachers need extended lengths of time to meet with colleagues and a consultant; meetings should be built around needs as perceived by the teacher; a specific goal should be designated; and teachers should understand the pedagogical processes in addition to skills in a stimulating context (pp. 70-72).

Shanker (1990) insists that the "locus for staff development is in the school" (p. 93) as a continuous, ongoing part of the school day, not something to be conducted after school hours or at another site. However, Joyce, Bennett and Rolheiser-Bennett (1990) contend that the location of the workshops or staff development is unimportant as "some highly effective programs have been run in all manner of settings" (p. 36).

What happens in central office is also crucial to the success of any staff development program. According to Fullan (with Stiegelbauer, 1991), "the district administrator is the single most important individual for setting the expectation and tone of the pattern of change within the local district" (p. 191) while studies have shown that "the chief district administrator and central district staff are an extremely important source of advocacy, support, and initiation of new programs" (p. 54). At the 1994 National Council on Measurement in Education conference, one of the themes which emerged from sessions and discussions was the
importance of administrative support and encouragement for teachers experimenting with alternative assessments (Conference highlights, 1994). Miller (1988) argues for beginning change in the central office, but involving teachers in all aspects of staff development even to the extent of utilizing the school's own teachers as experts. This sends a strong message to the teachers, "It proclaims that professional knowledge is not a commodity to be contracted from district and degreed consultants but can be found locally in classrooms throughout the district" (p. 174).

Miller contends that it is important for administrators to continue to view themselves as teachers as they move to central office:

Keeping this perspective requires an ongoing engagement with teachers and a conscious effort to understand and appreciate the complexity of teaching as a craft. Armed with this perspective, administrators have only to provide teachers with the opportunity to work together, the resources to do the job, and the latitude to work at their own pace and in their own style. (p. 183)

Fullan (with Stiegelbauer, 1991) concurs that "frequent, personal interaction is the key to implementation" (p. 199) because they see communication as being a big problem even for administrators who have thought of "everything" (p. 199). "To the extent that the information flow is accurate, the problems of implementation get identified. This means that each individual's personal perceptions and concerns - the core of change - get aired" (p. 199).
Other assumptions about staff development from Wood and Thompson (1993) deal with the role of the principal as "key in any effort to improve professional practice" (p. 54). They state:

Successful improvement programs usually find principals working collaboratively with their faculties to establish goals, working with teachers to plan staff development, participating in learning with teachers, providing follow-up assistance, and serving as an advocate for teachers who are implementing more effective professional practices in their classrooms. (p. 54)

Fullan (with Stiegelbauer, 1991) sees the principal as the middle manager sometimes caught "between the teachers and external ideas and people" (p. 144). They list two things that principals have done in effecting change:

They showed an active interest by spending time talking with teachers, planning, helping teachers get together, and being knowledgeable about what was happening. And they all figured out ways of reducing the amount of time spent on routine administrative matters; they made sure that change had an equal priority. (p. 168)

However, Fullan (1992a) contends that forceful principals who move into a school and try to implement radical change may not be effective over the long term. The change may occur to some degree but may disappear when the principal leaves. According to Fullan, the problem is that "too much store is placed in the leader as solution as compared to the leader as enabler of solutions" (p. 19). The principal should provide guidance and support, but also allow the teachers to express ideas and
leadership as well. The principal should also respect and value the teachers' wisdom and experience.

Another important assumption concerning staff development implies that knowledge about adult learners should serve as the guide for planning and implementing staff development (Wood and Thompson, 1993). Knowledge about adult learning includes ideas that:

Adults will learn, retain, and use what they perceive relevant to their personal and professional needs; adult learning is ego involved; asking an educator to learn and implement new professional behaviors may promote a more or less positive view of self. (p. 55)

Adults want to be the origin of their own learning and will resist learning situations that they believe are an attack on their competence. (p. 55)

One other relevant assumption was that, in addition to follow up, there needs to be some way to recognize and reward the implementation of the new practices.

As for the teacher's role in change and implementing innovations mandated or suggested from higher up, Fullan (with Stiegelbauer, 1991) summarizes the situation best: "Educational change depends on what teachers do and think - it's as simple and as complex as that," (p. 117) while it is also dependent on what those others (administrators, policy-makers, parents) around them do. Fullan, further adds:

Change is a highly personal experience - each and every one of the teachers who will be affected by change must have
the opportunity to work through this experience in a way in which the rewards at least equal the cost. The fact that those who advocate and develop changes get more rewards than costs, and those who are expected to implement them experience many more costs than rewards, goes a long way in explaining why the more things change, the more they remain the same. If the change works, the individual teacher gets little of the credit; if it doesn't, the teacher gets most of the blame. (p. 127)

Various roles are assumed by teachers during restructuring. According to Schlechty (1993), the trailblazers are those fearless leaders who are not afraid to take the first steps into the unknown. This role requires a clear guiding vision and those accepting this role must receive ample support from administration. They need flexible schedules so that they may visit other trailblazers and discuss what they learn from their visits. Those in this group also need acknowledgments for their efforts, such as opportunities to speak at conferences and seminars.

A second role that may be assumed by teachers during restructuring is that of the pioneer. Teachers in this group must also be risk takers, but their path has already been cleared for them by the trailblazers and there are fewer unknowns. However, their job is still difficult and requires much support and encouragement.

The next group to join a project is that of the settlers. This group needs guidance to know where they are going. They aren't risk takers like the ones that went before them and, therefore,
require strong leadership. They also need benchmarks and lots of feedback. Schlechty (1993) suggests that the best way to deal with the group known as "the stay at homes" is "benign neglect coupled with as much generosity as possible" (p. 50). Finally, there is the last group known as the saboteurs who try to stop change. Staff developers should understand each of these roles and the demands they require.

Mirel (1994) describes a reform project in which a district received $1.25 million dollars to create and implement a "break the mold" educational program that sprang from the concept of life-long learning and "implies that a whole community rather than just its schools must be the locus of any truly comprehensive educational reform movement" (p. 488). Despite adequate funding, administrative support, and a dedicated group of supporters, the project was dead within a year. A fundamental shift in philosophy that was required of educators in this project was "from teacher-centered to student-centered instruction and from school-based to community-based learning" (p. 491). Mirel claims that:

The Bensenville experience reveals that even the best-laid plans, no matter how well funded or how highly praised, will go awry unless reformers deal seriously and realistically with the issues of local control, school finance, the role of teachers and teachers' unions, and the hopes and fears of parents." (p. 515)
Teachers, parents, and students turned out to be the major critics of the project. Teachers "worried about what impact the restructuring would have on their jobs and the education of their students" (p. 495) while the parents and students were concerned with "changing the starting and ending times of school, lengthening the school year, and eliminating grades" (p. 495).

The general consensus seems to call for professional or staff development that acknowledges and works with the school culture, meets the needs of the teachers and the school, involves the teachers in the planning process, allows them to take some of the responsibility for their own learning, utilizes their knowledge and expertise in the instructional stage, while accomplishing this over an extended period of time rather than in one shot workshops. The role of the administrators and principal should be supportive but not overbearing, and finally, as pointed out by Joyce, Bennett, and Rolheiser-Bennett (1990) "poor training will not sustain enthusiasm among volunteers, just as good training will not overcome hostility at being forced to a training setting at gunpoint" (p. 35-36).

To accomplish all this requires restructuring and policy changes throughout the school system. Even though much is known about staff development, "it is still not well practiced" (Fullan, 1990, p. 3). Why? According to Fullan (1990), it takes wisdom, skill, and persistence to carry it out, and the other reason is political, "staff development is big business, as much related to
power, bureaucratic positioning, and territoriality as it is to helping teachers and students" (p. 4).

**Staff Development and Assessment Training**

Two themes which have prevailed throughout this review of the literature are (1) that assessment has come into focus as a key impetus in the nation's current educational reform movements, and (2) that teachers play a significant role in the success of any reform movements; they must see the importance of change and reform even before learning more about assessment in general, in addition to the alternative assessments that are currently being promoted.

In the area of assessment, Baron (1991) predicts that the focus of change rests with the teachers and "their access to rich assessment tasks and training in applying multidimensional scoring criteria" (p. 306). Worthen (1993a) also argues that the success of alternative assessment rests with the classroom teacher and his/her preparation to conduct it, while Stiggins (1991a, 1991b) warns that money must be made available for teacher training if we are to ensure quality assessments. Neill (1993) stresses the need for a socially responsible assessment system which is "centered primarily on the relationship between the teacher and the students" and which is "fair and equitable" (p. 24). He recommends that the district's role in this change is to offer support by: making class size smaller, providing more staff
development time as well as opportunities to work with other staff members. In addition, the district should "support the process of classroom-based assessment - by ensuring that teachers are trained in these new skills and by helping set, develop, analyze, and refine standards for all children," plus "analyze the assessment outcome data and present the information in a comprehensible format that is useful for improving the district and the schools" (p. 25).

Worthen (1993b) offers ten conditions of readiness to be considered before embarking on a move to alternative assessment. The conditions are:

1) desire for better assessment information,
2) indications that current assessment is creating negative side effects,
3) staff openness to innovations,
4) conceptual clarity about assessment,
5) alternative assessment, assessment 'literacy',
6) clarity about desired student outcomes,
7) content or curricula ill-suited to traditional tests,
8) school examples of alternative assessment,
9) willingness to critique assessment methods,
10) patrons' and policy makers' openness to new forms of assessment. (p. 455-456)

Authentic assessment is consistent with a constructivist paradigm, and so other aspects of instruction, curricula, as well as grading, and other school policies must be consistent with that paradigm, if it is to be used. Martin-Kniep and Kniep (1992) argue that "rather than tinkering around the edges of the existing system or attempting to make piecemeal changes in programs or
schools, the process (which evolved from Education 2000) is to result in a system-wide infrastructure that is ultimately enabling of redesign at the school and classroom levels" (p. 7). They envision soliciting the support and ideas of the entire local community.

Johnson (1992), goes on to suggest that the whole school must work together, starting from graduation and working backwards:

If a school has clearly defined what its graduates will know and be able to do upon graduating - if, in fact, they have developed a well-defined graduation exhibition that would show this - then the classroom teacher has clear guidelines to show the way. (p. 39)

According to Sperling and Mahalak (1993), a well designed workshop can be used to teach performance assessment to teachers. Their workshop includes six components: assessment theory, a performance assessment model, teachers design classroom assessments, the role of students using performance assessment, student self-assessment impacts on teacher practice, and workshop evaluation and future plans. The success of the program is attributed to these factors: limited workshop membership, four days, four hours per day, ample time to learn and digest information, two instructors which allowed for feedback and conferencing, administrative support, hands on learning, individual teacher needs met, and three, two-hour, after school follow-up sessions.
As with any new trend, enthusiasm runs high, as many hope for solutions to all our present problems. Although performance, authentic, or alternative assessments address many of the problems, nonetheless, these tests are too new to have acquired much of a track record; as Haney and Madaus (1989) point out, we need to adopt a "critical attitude." They suggest that these alternative assessments need to be closely monitored and research needs to be conducted before their effectiveness can be fully known. Worthen (1993a) suggests that these assessments must be given time to prove their success or failure. He warns that results and solutions to all educational ills are not possible and should not be expected. Overall change requires more than just redesigning tests, it requires the reeducation of so many: classroom teachers, teacher educators, administrators, instructional designers, politicians, parents, and students (Stiggins, 1991a), as well as a change in attitudes and instructional practices. And the key to accomplishing all of these begins with research and more in-depth studies.

Preliminary results of studies on authentic or performance assessments are beginning to be shared at conferences and in journals. CRESST (Sheperd and Bliem, 1993) is conducting studies as part of a larger project to "examine the effects of implementing performance assessments on instruction and student learning" in mathematics and reading. Researchers from the University of Colorado at Boulder worked with a school district outside Denver
and specifically with third grade teachers from three schools. Flexer and Gerstner's study (1993) reported on issues and concerns of teachers developing performance assessments in mathematics. The researchers were to help teachers design performance assessments for mathematics and to support implementation of changes in instruction and assessment. The initial issues encountered by teachers concerned (1) organizational support, and (2) lack of time. In an effort to address these, researchers allotted extra time in the form of course credit and provided the teachers with resources while the district allowed a half day a month for work on the project. The major issues which ultimately surfaced had to do with "beliefs and practical teaching knowledge. For some of the teachers the real struggle was not in learning how to use performance assessment, but in believing that it was a useful thing to do" (p. 30).

Hiebert and Davinroy (1993), also working on the same project, looked at issues "that arise over the course of staff development on classroom-based assessments" (p. 4). During the project, a moratorium was called on the standardized tests for the third grade of the participating schools. This allowed the teachers in the study to choose and adapt assessments to their classrooms without having the worry of preparing the students for the standardized tests. During the initial sessions, the research team allowed the schools and teachers to identify their particular goals and appropriate assessments, and then "subsequent sessions
involved developing scoring rubrics among participating teachers" (p. 5). One dilemma encountered was the question of how the assessments fit with their instructional programs (p. 11).

Another issue which arose was change in instructional practices due to the assessments. A dilemma which surfaced had to do with the incompatibility of the assessments with the school's philosophy. They also encountered problems with sharing information about the students with the parents. However, as they reflected and struggled with these issues, they decided to make changes in their instruction and standards based upon the assessments. This group of researchers caution staff developers and school administrators who wish to implement alternative assessments that "the responses of teachers in these projects may vary greatly" and "that the process is not necessarily one welcomed by teachers" (p. 24).

Sheperd (1995) reports on these studies after the second year. The major issues that arose during the project had to do with differences in views of researchers and teachers, and time. The researchers assumed that all the teachers were true volunteers, but that was not exactly the case.

The time issue has several components. The teachers felt that they did not have enough time to do the planning, grading, meeting with teams, and simultaneously doing things the old ways while trying new things. These complaints were handled in a couple of ways. The district offered support by granting a half-
day release time per month and allowing the teachers to use district inservice time to work on the new activities.

There were several positive outcomes from the project. First, "teachers involved their students in the development of scoring criteria and in grading one another's summaries" (Shepard, 1995, p. 41). In addition, the "teachers eventually developed greater sophistication about scoring criteria" (p. 41). Most importantly, there were some hints for future staff development programs. Shepard emphasizes the importance of having the teachers understand "the philosophical and conceptual bases of the intended curricular goals" (p. 42). Further, the teachers need:

1. appropriate materials to try out and adapt;
2. time to reflect and to develop new instructional approaches;
3. ongoing support from experts to learn (and challenge) the conceptual bases behind intended reforms (p. 42).

Shepard enthusiastically reports the positive outcomes of their project on using performance assessments for reforming education, but offers this caution:

If teachers are being asked to make fundamental changes in what they teach and how they teach it, then they need sustained support to try out new practices, learn the new theory, and make it their own. (p. 43)

Further findings were shared by researchers at the 1995 annual meeting of the American Educational Research Association. Speaking at the conference, Borko (1995) discussed the
the importance of considering teachers' beliefs. She cautioned that if their beliefs are incompatible with the intent of the staff developers and are not addressed, the teachers are likely to ignore new ideas. Putnam (1995) responding to all the papers presented at the session commented on the role of teachers' knowledge and beliefs effecting change and the importance of both sides - teachers and those offering the new ideas - to listen to each other, and what it means to collaborate when doing research.

Research has shown that parents play a huge role in the success or failure of new programs in schools. This holds true for the introduction of authentic assessment as well. Shepard and Bliem (1993, 1995) report on the importance of communicating with parents when introducing authentic assessments into a classroom or system. Working within the same University of Colorado project mentioned earlier, Shepard and Bliem conducted a study concerning parent's views of authentic assessments. Parents were shown examples of both authentic assessments and traditional assessments and then asked questions. Their responses showed approval "of both types of measures, giving stronger approval ratings to performance assessments" (1995, p. 31).

Results of a three year study sponsored by the Office of Educational Research and Improvement reveal the affect of performance assessments on teaching and learning (Kane and
Khattri, 1995; Khattri, Kane, and Reeve, 1995). The study shows that "the effect of assessments on the curriculum teachers use in their classrooms has been marginal, although the impact on instruction and on teacher roles in some cases has been substantial" (p. 80). The researchers concluded that factors effecting the success or failure of assessment reform have to do with administrative support, including opportunities to experiment, professional development, and informal assistance to the teachers' implementation of the assessments (Khattri, Kane, and Reeve, 1995) as well as the "coordination between purposes and assessment formats, connections between assessment and curriculum, and cost" (Kane and Khattri, 1995).

**Chapter summary**

This chapter has presented a descriptive account of several areas of research which are relevant to this study: learning, testing, assessment training, change and staff development, and studies on authentic assessment.
CHAPTER 3
METHODOLOGY

Introduction

This chapter presents the inquiry methods used to conduct this case study of the introduction and implementation of authentic assessment in a school and classroom. The chapter is divided into three sections. The first section describes the findings of the pilot study, the rationale for qualitative methodology, and the site and population of the study. The second section explains the data collection process including the methods used to collect data, the amount of time spent at the site, the documents that were collected, and the researcher's role as instrument. The third section discusses the data analysis process and how the researcher established trustworthiness through credibility, transferability, confirmability, and dependability.

SECTION ONE
Preliminary Fieldwork

To develop more fully my research questions and inform my research design, two pilot studies were conducted. During the 1992-1993 school year, I observed a teacher who utilized
authentic assessment practices in a school where such practices are encouraged; the second pilot was conducted during the first several weeks of the 1993-1994 school year at the site of the proposed study. I entered the pilots with an open mind, content to take in what was happening in the classrooms in the hopes that specific questions would emerge or, as Bogdan and Biklen (1992) suggest, "to mentally cleanse" and to allow "plans [to] evolve" (p. 58).

In the interviews of students and teacher in the first pilot, issues of time constraints, grading, learning, risk taking, motivation, and administrative support were raised. The teacher was overwhelmed by the time and energy it takes to create, implement, and grade authentic assessments, although she could see benefits in terms of student learning and understanding.

The early observations in the proposed site yielded information on the difficulties of implementing changes and programs for and with teachers due to their busy schedules and previous commitments. An assessment course for the teachers which was to begin in September had to be postponed because some teachers had previously committed themselves to other courses or needed the time to meet pressures and expectations of them from new district and school programs and changes. Observations of several classrooms revealed initial trials of authentic assessments by two teachers who had learned about it.
from professional journals and a summer assessment workshop attended by district representatives.

**Design of Study**

The purpose of this study is to describe the introduction and implementation of authentic assessment into a school system. This necessitates considering not only the problems and issues raised during the staff development sessions introducing authentic assessment, but also the attitudes and concerns of teachers, students, and administrators regarding authentic assessment in the classroom. Teachers are the primary focus, but they cannot be considered alone since their actions affect and are affected by the attitudes and actions of others (students, administrators, and parents). In order to understand and discern the attitudes and perspectives of participants, the researcher cannot merely administer attitude surveys (Patton, 1990, p. 47), but must find out what people believe, think, and feel by getting to know them, then spending time with them, and participating in their lives, that is, doing fieldwork. This research addresses the questions: What is happening here? And, what are the local meanings of these happenings for those involved? (Erickson, 1986). Understanding requires qualitative methodology. According to Bogdan and Biklen (1992), "the research questions ... are formulated to investigate topics in all their complexity, in context" (p. 2).
Patton (1990) also maintains that there are many aspects and attributes of programs which cannot be quantified, but rather require a qualitative approach to deal with questions of "What do programs mean to participants? What is the quality of their experiences? Answers to such questions require detailed, in-depth, and holistic descriptions that represent people in their own terms and that get close enough to the situation being studied to understand firsthand the nuances of quality" (p. 109-110).

"A case study is a detailed examination of one setting, or a single subject, a single depository of documents, or one particular event" (Bogdan and Biklen, 1992, p. 62). It allows the researcher to start by casting "a wide net trying to judge the feasibility of the site or data source for their purposes" (p. 62). Choosing a focus for the study, whether a school within a district or a teacher within a school, "is always an artificial act, for you break off a piece of the world that is normally integrated" (Bogdan and Biklen, 1992, p. 63), but is a necessary step in order to "make the research manageable" (p. 63). This is certainly the case with this particular study. Aware of the need to consider all aspects of the community; the school; the teachers involved; the students; the administrators; the assessment course; the senior seminar course; and the teacher teams; I had to consider the feasibility of one person looking at all these components and single handedly managing to collect and analyze the data. I then had to make
decisions as to what should be included and what should be dismissed or ignored as less relevant.

Alternative assessment is, itself, qualitative rather than quantitative in nature. The goal of such assessments is to produce richer, clearer pictures of the knowledge the students have constructed rather than to produce numerical measurements for comparison purposes. Studying such qualitative assessments requires qualitative methodology to capture a more informed understanding of how the teachers view such assessment. This type of assessment works in concert with instruction and is based on a theory of learning which reflects a constructivist epistemology. In addition, as Glesne & Peshkin (1992) state, it is important for the method to match one's world view. I believe that one's epistemological philosophy is intertwined with the kinds of teaching, assessment, evaluating and research one employs, and enter this study subscribing to a constructivist epistemology.

Population and Setting

The site of the study was a small, rural school district in New Albany, Ohio with an enrollment of about 800 students. The district was selected for specific reasons. In the last few years, developers had purchased land around the township with the idea of building a community. Construction of homes soon followed, and changes were made in the school district including
appointment of a new superintendent who promoted the use of authentic assessment. In the fall of 1993, an assessment course was planned for the teachers. The district, therefore, offered a purposeful sample of classrooms where teachers were receiving training in assessment and where authentic assessment was being encouraged and supported. Purposeful sampling was used to insure "information-rich cases for study in depth" (Patton, 1990, p. 169) to yield data regarding issues appropriate to the purposes of the study. The district administrators were also interested in working with The Ohio State University by offering the schools' classrooms as work sites for students and researchers. In August, 1993, I learned of the district's intention to offer a staff development assessment course. Therefore, I used my role as an Ohio State student/researcher to gain entree into this school district (Burgess, 1991).

This district offered an uncommon opportunity for me to observe the implementation of the most recent educational theories and ideas being put into practice. However, Marzano (1991) cautioned that in order for a new innovation to work, it must not contradict the system's existing paradigms. "If it does, then the innovation must be accompanied by a paradigm shift - the introduction of a new paradigm that supports the innovation" (p. 411). This meant that introducing a new idea, specifically, authentic assessment into the system was not enough, the idea needed to be compatible with the existing epistemological
philosophy or accompany an epistemological shift consistent with constructivism. It was (and is) my belief that the leaders of this community have been working toward that shift.

I began my observations at the school in August, at the start of the 1993-94 school year, because "early-year encounters between teacher and students are of special significance. Unusual, unrepeated and important things happen in these initial encounters" (Ball, 1990, p. 163), and many of those encounters have to do with the guidelines and criteria for the year's assessment. I attended the initial staff meeting, the day before school began, and was introduced to the teachers and the rest of the staff at that time. Discussions that spring and over the summer between district administrators and The Ohio State University faculty had resulted in plans for an assessment course to be offered to district teachers in the fall quarter. Subsequently, the course had to be postponed until January. So, during the autumn months, I observed in classrooms throughout the high school and spent time in the teachers' lounge at lunchtime talking with the teachers and the principal in order to more fully develop my research questions and to narrow the focus of my study.

When the assessment course started in January, I attended the classes as a participant observer. Eventually, my focus narrowed to the senior team which included the four teachers who taught the Senior Seminar course. I had visited these teachers' classrooms several times during the fall and was
impressed with their enthusiasm, knowledge, and dedication to their respective disciplines and to the course that they had developed and were piloting that year. I soon felt that I had developed a good rapport with them and that we could benefit from working with one another. The major problem I encountered was my reluctance to impose upon these teachers who were already inundated with so much work in their courses and the additional roles they assumed within the school. All four of the teachers were chairs of their respective departments and taught advanced placement courses, in addition to their regular classes. Originally, I observed these teachers in their classes: math, English, social studies, and science. However, as the assessment course developed and the teachers were asked to devise philosophies for their assessments and devise new assessment plans, these teachers decided to work together to create a philosophy and plans for the Senior Seminar pilot rather than their individual courses. So, several weeks into the assessment course, my focus moved to these teachers and their work on the assessments for the Senior Seminar course.

SECTION TWO
Data Collection

The observations and interviews from the pilot studies guided my early observations in the field. Figure 3.1 shows a timeline of the study.
<table>
<thead>
<tr>
<th>COMMUNITY SCHOOL RESEARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
</tr>
<tr>
<td>March</td>
</tr>
<tr>
<td>Trouble passing levies, 2nd &amp; 3rd Generation, Why Change?</td>
</tr>
<tr>
<td>May</td>
</tr>
<tr>
<td>High School ranked worst in county by <em>Columbus Monthly</em></td>
</tr>
<tr>
<td>July</td>
</tr>
<tr>
<td>Sept.</td>
</tr>
<tr>
<td>New School Board</td>
</tr>
<tr>
<td>Nov.</td>
</tr>
<tr>
<td>Superintendent Quits</td>
</tr>
<tr>
<td>New Superintendent</td>
</tr>
<tr>
<td>1993</td>
</tr>
<tr>
<td>March</td>
</tr>
<tr>
<td>New School Board</td>
</tr>
<tr>
<td>May</td>
</tr>
<tr>
<td>Superintendent Quits</td>
</tr>
<tr>
<td>New Superintendent</td>
</tr>
<tr>
<td>July</td>
</tr>
<tr>
<td>Sept.</td>
</tr>
<tr>
<td>Groundbreaking for new school</td>
</tr>
<tr>
<td>Nov.</td>
</tr>
<tr>
<td>Wiggins' Workshop</td>
</tr>
<tr>
<td>1994</td>
</tr>
<tr>
<td>March</td>
</tr>
<tr>
<td>Construction Halted on new school</td>
</tr>
<tr>
<td>May</td>
</tr>
<tr>
<td>Conflicts between Township, Village and Columbus</td>
</tr>
<tr>
<td>July</td>
</tr>
<tr>
<td>Sept.</td>
</tr>
<tr>
<td>Passed School Levy</td>
</tr>
<tr>
<td>Merger didn't pass</td>
</tr>
<tr>
<td>Nov.</td>
</tr>
<tr>
<td>1995</td>
</tr>
<tr>
<td>Construction Resumed</td>
</tr>
<tr>
<td>March</td>
</tr>
<tr>
<td>May</td>
</tr>
<tr>
<td>2nd Exhibition</td>
</tr>
</tbody>
</table>

![Timeline of the study](image)

*Figure 3.1: Timeline of the study*
Prolonged engagement of two school years and interaction in several classes allowed me to feel "the everyday experiences of subjects" and to establish rapport (Glesne & Peshkin, 1992) with the participants. My goal was to conduct my research "with, not on" (McLaren 1991, p.154; Noffke 1990, p. 1) those involved. Erickson (1986) noted that "it is necessary to begin observing and interviewing in the most comprehensive fashion possible. Later in the research process one moves in successive stages to more restricted observation focus" (p. 143). As mentioned above, I began my research by observing several high school classrooms looking at the assessment employed there and the reactions of students and teachers to the assessments. I moved my focus to the teachers involved with the assessment course, and eventually narrowed my focus to the senior seminar team. Although my observations dealt mainly with what was happening in the classrooms, I also attended and observed other functions such as staff meetings, inservice meetings, planning sessions for Senior Seminar, parents' night, exhibition nights, as well as informal gatherings such as occurred at lunch time.

The assessment course that was offered to the teachers met for two and one half hours from 3:30 to 6:00, six times during winter quarter from January through the beginning of March. Thirty teachers, three administrators, and the guidance counselor signed up for the course. The participants were to receive two graduate credits for attending the course. The professor teaching
the course was a full professor with 22 years experience teaching, researching and designing assessment at The Ohio State University.

I attended the assessment course sessions, taking notes on the content of the course, observing the teachers' reactions and talking with them about assessment and their thoughts and feelings concerning the course as well as what they were learning. I also offered my assistance to the teachers as they worked on course assignments which included designing new assessments and rubrics for their classes. My intention going into the study was to have the teachers keep journals to document thoughts, ideas, and questions prompted by the course, classroom assessment, or their interactions with me, but it became evident that the teachers were already so overloaded with work from their classes, other expectations, and the course, that this seemed an excessive request, so I asked them to contact me and to let me know of any problems or concerns with the course or their implementation of the assessments. Several teachers did send e-mail messages, while others talked with me in the teachers' lounge or after class. I wanted my role to go beyond being a passive observer; I tried to be a participant observer working more as a collaborator with the teachers as modeled in Hunsaker and Johnston (1992) and Noffke (1990).
Interviews

The interview is "an interaction between two human beings" (Glesne & Peshkin 1992, p. 83) and its purpose is to inquire about those things which cannot be seen, that is, attitudes and feelings of the participants about the topic (Glesne & Peshkin). The interview literally allows an "inter-view" as it is the "interchange of views between two persons conversing about a theme of common interest" (Kvale 1992, p. 1).

In the early part of the study, I interviewed the teachers who comprised the senior team concerning their attitudes, actions, and comments as interpreted from my observations, from the pilot study interview, and from readings of the literature. Each interview lasted about 50 minutes. I began with several structured questions: How long have you been teaching? How long have you been at New Albany? What do you see as the purpose of assessment? What kinds of assessments do you use? What does the term authentic assessment mean to you? Do you use any kind of assessment that you consider authentic? Why are you taking the assessment course? What do you hope to gain from the course? New questions emerged during the interviews, and I encouraged the teachers to elaborate on my questions and to add their untapped concerns as well. With the participants' permission, the interviews were audio-taped, and later transcribed verbatim. Group interviews of the senior team were then conducted at the end of the school year, again in December,
1994, and finally in May, 1995 to document the changes in attitudes and practices which occurred during and after the teachers' participation in the assessment course and the implementation of the assessments in the second year of the course.

The data from the first set of interviews along with remarks made by the teachers and administrators, and questions I encountered as the study developed, were used to construct two paper and pencil, grounded surveys, the first of which was administered to all teachers attending the assessment course in January as the course began, and the second one in March, at the conclusion of the course. The grounded survey grew out of my observations, interviews, and interactions. As I reviewed the collected data and began to "see" emerging themes, I created a questionnaire as a means of checking and perhaps confirming my hunches. As the data led in certain directions, I used this information to develop the grounded survey to further establish a "grounded theory" as outlined by Bogdan & Biklen (1992). The surveys of the teachers taking the course helped to corroborate my observations, interviews, and analysis of documents.

At the end of the 1993-1994 and the 1994-1995 school years, I also interviewed several students from the Senior Seminar, regarding their thoughts and reactions to the assessments used in Senior Seminar throughout the year and how they compared with more objective testing. The students that I
interviewed represented those who had received awards or honorable mentions for their projects as well as those who had done very little for their projects and received low marks. An exit survey was also administered to all seniors at the end of each of the school years.

In many cases, administrators set the tone for the instructional environment of a school which also includes the forms of assessment which are mandated, encouraged, permitted and/or rejected. When these alternative assessments have been introduced, one of the critical issues is the extent and type of support that is provided by the administration (Flexer & Gerstner, 1993). In addition to administrators' personal attitudes about assessment, the district may have certain policies concerning the types of assessment to be used, the frequency of certain tests, and subsequent grades.

Therefore, I interviewed three administrators from the school who have the authority or duty to effect the assessment used within the district and individual schools: the superintendent, the administrator in charge of curriculum and instruction, and the high school principal. These interviews were also audio taped and transcribed.

All the interviews and observations were given codes so that quotes taken from them could be traced back to their origin. The first letter referred to the person or team, for example, D for Mrs. Dearheart, J for Dr. Jones, T for team, S for student. The
second letter symbolized its source, whether it was from an interview or observation (I or O). The first number indicated which interview (1 for first, 2 for second), and then the date of the interview or observation (5-95 meant May, 1995).

I met several times with the instructor of the assessment course to discuss the format, participation, progress, and any other pertinent information regarding the course. I also conveyed to him many of the expectations and needs of the teachers as reported to me by them.

Observations of Senior Seminar included the presentations of the final student projects, and the culminating exhibition for the general public in both 1994 and 1995. The judges for the exhibitions had been selected from the community at large. Parents, siblings, peers, teachers, and a few other community members attended.

After the 1993-1994 school year ended, the district held a three day inservice for the teachers which I also attended. The senior team worked on redesigning the course and preparing the graded course of study to present to the school board in July to make Senior Seminar a graded course as well as a requirement for graduation.

**Document Analysis**

In addition to the interviews, field notes recorded during observations, and the grounded surveys, I collected numerous
documents, including, but not limited to, the strategic plan for the district, grant proposals, scores for the district's standardized tests for the last eight years, newspaper articles about the district and the Senior Seminar course, the assessment philosophies and rubrics created by the teachers for the assessment course assignments, the assessments created by the teachers doing subsequent independent study with the professor, a survey done by the district, as well as samples of the seniors' portfolios, journals, and a video of the projects presented exhibition night. Bogdan and Biklen (1992) have suggested that documents can be used to "corroborate your observations and interviews and thus make your findings trustworthy" (p. 52). The subjective nature of alternative assessments has been criticized for almost a century (see Starch and Elliot, 1912); concern with subjectivity played a major role in the promotion and growth of objective testing. Therefore, it was necessary to pay close attention to the criteria and rubrics the teachers were using to grade these alternative forms and the consistency with which they graded them.

**Researcher's Role as Instrument**

"As ethnographic researchers, we actively construct and are constructed by the discourses we embody and the metaphors we enact. We are, in effect, both the subject and the object of our research" (McLaren, 1992, p. 80). As a way of tracking and monitoring my subjectivity, I documented my thoughts, hunches,
and opinions, as well as observations in my reflexive journal. The specific topic chosen for study was based on my interest in authentic assessment, and at the outset of this study, I acknowledged my bias towards the positive effects of authentic assessment and my belief that teachers would gain confidence and knowledge concerning their abilities to assess their students' learning from the assessment course. I realize that I am the filter through which all the observations, interviews, and document analysis must flow. I determine which participants to involve, what questions will be asked, which classes will be observed and everything else that flows through me as filter to become the written story. According to Glesne and Peshkin (1992), "my subjectivity is the basis for the story that I am able to tell. It is a strength on which I build. It makes me who I am as a person and as a researcher, equipping me with the perspectives and insights that shape all that I do as a researcher, from the selection of topic clear through to the emphases I make in my writing" (p. 104).

SECTION THREE

Data Analysis

"Data analysis involves organizing what you have seen, heard, and read so that you can make sense of what you have learned" (Glesne and Peshkin, 1992, p. 127). "As you work with data, you must remain open to new perspectives, new thoughts" (p. 128). Decisions of which data "to code, which to pull out, which
patterns summarize a number of chunks, what the evolving story is, are all analytic choices" (Miles and Huberman, 1984, p. 21). Data analysis was an ongoing activity throughout the study and so helped "focus and shape the study as it proceed[ed]" (Glesne and Peshkin, 1992, p. 127).

After each observation, or interview, and after each day in the field, I entered data into a word processor. While the computer sufficed to record the data, I attempted to code and identify themes and patterns by carefully pouring over fieldnotes, journal entries, documents, interviews and survey findings. "Coding represents the operations by which data are broken down, conceptualized, and put back together in new ways" (Glesne and Peshkin, 1992, p. 133). Several phenomena for coding surfaced during the preliminary fieldwork, for example, teacher's attitude toward assessment in general, teacher's attitude toward authentic assessment, teacher's purposes of assessment, teacher's use of authentic assessment, teacher's concern with authentic assessment, and various kinds of assessment used. I used these as a beginning point for identifying codes in the subsequent data. New codes were added as they emerged throughout the study. These codes had to do with the perceptions, attitudes, concerns, practices, and actions of the students, administrators, and teachers regarding alternative assessments and the course.

Initially, descriptive code names were assigned to data and notes were written in the margins of transcribed interviews,
observations, field notes, and surveys. Later in the analysis, the data were coded according to themes and patterns in order "to understand the patterns, the recurrences, the whys" (Miles and Huberman, 1984, p. 67).

Displaying data may be useful to the researcher to "eyeball data; carry out detailed analyses; set data up for another, more differentiated display; combine parallel data for a single site; or combine data from several sites; report findings" (Miles and Huberman, 1984, p. 79). I created a context chart and a role-ordered matrix to combine parallel data from the site (Miles and Huberman, 1984).

The grounded surveys were also searched and coded for similarities and differences to establish common concerns and beliefs and isolated concerns and beliefs among the staff. However, I expected that discrepancies and contradictions would appear in the data. I, therefore, was not looking for "the truth" as in the positivist paradigm, but for a perspective on "what does this person's perception mean and how does it effect his/her actions". So I considered all respondents' statements valid to some degree, and felt they needed to be considered from their perspective and within context. As Kvale's quotation of the Thomas theorem states, "if men believe ideas are real, they are real in their consequences" (Kvale 1989, p. 81).
Data Story

One of the most important aspects of the study is to share the findings with some specified audience. Some researchers (Denzin, 1994; Richardson, 1994; Wolcott, 1990) have lamented that many qualitative studies are so sterile, dull, and inaccessible that reading through the findings to the conclusion becomes burdensome or difficult. I wanted my research to serve those in areas beyond academia; I wanted it accessible to teachers, administrators, and parents, as well as the general public. Therefore, I have tried to write in a less academic and a more general format.

Moreover, informed by the feminist and post modern literature, I recognized that the findings or data which I presented were only a small slice of the whole and offered merely a partial story constructed by me. According to Wolcott (1990), this has been both my challenge and my opportunity, "to portray real people doing real things through the eyes of another human observer" (p. 49). Therefore, in trying to write up my findings, I gave the participants as much voice as possible by using their own words from their interview's whenever appropriate, and in some instances I tried to create more interesting and readable data stories by utilizing experimental formats as exemplified by Lather (1991), Merryfield (1992), Patai (1988), Richardson (1994), and VanMaanen (1988).
Trustworthiness

Lincoln and Guba (1989) offer guidelines to assure rigor in qualitative research. They suggest establishing trustworthiness by substituting different criteria for those used in quantitative research, that is, reliability, internal and external validity, and objectivity. The parallel criteria of credibility, transferability, confirmability, and dependability are more meaningful within a constructivist epistemology.

Credibility

Credibility is the analogue to internal validity, but instead of comparing the data to reality, it is looked at from the point of view of the respondent's interpretation or construction of reality against the reality as presented by the researcher. This is done by staying longer and getting to know, discussing with uninvolved persons, accounting for all known cases, monitoring one's reflexivity, and confirming data with those from whom it was collected.

Prolonged engagement and persistent observation in the field allowed me to collect more data, but also allowed me time to develop trust and rapport with the participants. My research was conducted over a period of 21 months. Progressive subjectivity or "the process of monitoring the evaluator's own developing construction" (Lincoln & Guba 1989, p. 238) was accomplished by maintaining a field log and a reflexive journal in
which I recorded my observations, thoughts, hunches, concerns, and reflections on what I saw, heard, and discussed. This continued throughout the study as I conducted interviews and surveys.

My findings are interpretations of what I saw and heard, and these were offered back to the respondents as member checks to authenticate my construction of their world. Following the interviews, I gave each of the teachers a transcribed copy of his/her interview and asked them to read it to confirm its accuracy. Later, following data analysis, the teachers and administrators were given a copy of the final report. They were requested to check it for inaccuracies and make comments on my interpretations of the data. "The field researcher needs to share with his or her subjects the discourses at work that are shaping the field site analysis, and how the researcher's own personal and intellectual biography is contributing to the process of analysis" (McLaren, 1991, p. 158).

Using negative case analysis I tried to account for all known cases and attempted to adjust my hunches accordingly, but I was aware that probably not all cases would fit into the categories created.

"One important way to strengthen a study design is through triangulation, or the combination of methodologies in the study of the same phenomena or programs" (Patton, 1990, p. 187). Triangulation of sources was employed to check out "the
consistency of different data sources within the same method" (Patton, 1990, p. 464) by (1) comparing the findings of the observations, the interviews, the surveys, and the numerous documents collected throughout the study, and (2) "comparing the perspectives of people from different points of view" (Patton, 1990, p. 467). I was cautious about expecting all the sources to lead to a totally consistent picture, however, based on Patton's claim "that the different kinds of data have captured different things and so the analyst attempts to understand the reasons for the differences. At the same time, consistency in overall patterns of data from different sources and reasonable explanations for differences in data from divergent sources contribute significantly to the overall credibility of findings" (p. 467-468).

As another means of verifying and validating my data and interpretations, I also used analyst triangulation (Patton, 1990, p. 464). As part of the ongoing study, I conversed with two doctoral students and we offered each other unbiased viewpoints and detached perspectives as we discussed and read each others' accounts. These peer debriefers who were not involved with my research, but with whom I could discuss the data and findings, provided insight to the study. One debriefer had taken the qualitative research course with me and we discussed the data in terms of our mutual readings.
Transferability

Transferability is the parallel to external validity or generalizability. An exact replication of a qualitative study is not possible because of its emergent design. The responsibility for deciding whether the results could be transferred to another situation resides with the reader, not with the researcher/writer. No claims are made beyond that this is a study of a particular school, of these particular students, teachers, and administrators at a certain date in time; in other words, this is a slice of life, nothing more. The prime technique for establishing transferability is thick description (Lincoln & Guba 1989). My research process is well documented with as many details as possible, "of the time, the place, the context, the culture" (Lincoln & Guba 1989, p. 241) as they relate to the study, so that the reader of the study may decide whether the subjects, environment, and other variables match well enough that the findings of this study are transferable.

Dependability and Confirmability

Dependability is the parallel of reliability in the positivist paradigm. Confirmability is the parallel to traditional research's objectivity and both are addressed through an audit trail in which I describe my path of collection, analysis, and reduction of data. The written field notes, transcribed interviews, and data analyses document the process of the study. In addition, my reflexive
journal or field log tracks the visits, the events, and my hunches and insights as they occurred throughout the study (Bogdan and Biklen, 1992; Glesne and Peshkin, 1992) in an effort to document what and when evidence appeared which disconfirmed my preconceived notions, and to show "specific ways in which the changes in interpretive perspective took place" (Erickson 1986, p. 152). Erickson recommends writing "a first-person account of the evolution of inquiry before, during, and after fieldwork" (p. 152) which I also did.

**Ethics**

Whenever one deals with others, ethical treatment becomes a necessary concern (Eichelberger, 1989). I addressed this issue by first submitting an abstract of my study to the Human Subjects Review Committee. Throughout the study, I tried to maintain my responsibility to the rights of the participants. Administrators, students, and teachers were advised of the purpose and questions as they pertained to the study, and were asked to sign consent forms which stated that they could withdraw from the study at any time. During the study, I tried to be sensitive, understanding, respectful, thoughtful, and appreciative of the participants' time, comments, concerns, and I tried to give back to them by being an empathetic listener to their frustrations and stories regarding assessment, as well as offering my help in matters of assessment when appropriate. Confidentiality of participants' words and
actions was maintained by using pseudonyms for anonymity and being careful with whom I discussed these matters. However, because of the highly publicized nature of the community and school district in which I conducted my research, it was agreed that anonymity would be impossible, so permission was granted by the superintendent to use the name of the district. I maintained a close relationship with my advisor, who acted as the auditor of the study throughout the research, and I consulted with her whenever any questions or problems arose.

I also tried to avoid what McLaren (1992) describes as "absorbing, accommodating, homogenizing, or integrating (difference) into the totalizing schemes of Western essentialist ethnographic practices" (p. 90), and instead, attempted to "recover the meaning of identity as a form of cultural struggle, as a site of remapping and remaking historical agency within a praxis of liberation" (p. 90). There is such controversy surrounding assessment today that it seemed only "natural" to allow the teachers, students, and administrators to be a part of this debate, and as such, it became my role as ethnographic researcher "to act with the oppressed, not over them or on behalf of them" (p. 89).
CHAPTER 4
CONTEXT AND SEQUENCE

This chapter presents the context and sequence of events relevant to this study. Because of the uniqueness of the community and of many of the changes within that community and its schools, it is necessary to describe the following as they relate to this study: the village, the township, and the political climate of the school district; the high school and the Senior Seminar course, as well as the participants.

Until recently, the New Albany/Plain Local School District which serves Plain township and the village of New Albany was considered a rural, agricultural district. Recent upscale development has changed drastically the demeanor of this quiet, no-nonsense village and surrounding area as well as the sentiment of the residents; it is now described as a rural, suburban district.

As plans for the community evolved through the efforts of two businessmen and the New Albany Company, ideas for a world-class school emerged and were set into motion. I will describe that evolution and the persons who were instrumental in
it. Since the high school is the primary focus of the study, attention will be given to the present facility and its principal.

The Senior Seminar course is the result of ideas conceived by the superintendent, high school principal and a social studies teacher. Several schools throughout the United States require exit exhibitions, senior theses, or other programs for their seniors, but none have incorporated all the components of the New Albany Senior Seminar. Therefore, an explanation of the course and its evolving program and assessments is also necessary.

The second section of this chapter will document the sequence of events which led to the implementation of authentic assessment into the district and the individual classrooms, specifically, the Senior Seminar course. Direct intervention by the administration through a workshop, an assessment course, and inservice sessions will be outlined as well as the informal means of teacher access to authentic assessment.

CONTEXT

Community

In the fall of 1993, when research for this study began, the trip to the study site was reminiscent of a pleasant Sunday afternoon drive in the country. The busy freeway leading from the midwestern metropolitan city with its high rises and hustle and bustle gave way to a narrow road winding past corn fields to the old mill-general store. The village itself was redolent of an
era seldom seen today with its older homes, gun store, Dairy Mart, pizza parlor, barber shop, abandoned gas station, and the cinder block Municipal Building, all nestled together. A drive along the same road two years later reveals a different scene: vast areas of road construction, new housing, miles of white fences, huge concrete pipes lying along the road, the large construction site of the new middle school and high school campus, and some newly remodeled brick buildings standing alongside the older establishments in the village. The contrasts of old and new reveal a community in the midst of transition, but belie the tension and dissonance within and among the residents, both old and young. The three school buildings just a stone's throw away from the mill and volunteer fire station attest to the small village atmosphere.

The village of New Albany was founded in 1837 and as of 1990 had a population of 1,621 and household size of 2.73. The population of Plain township which includes New Albany was 4,355. In the mid 1980s, a well known, wealthy businessman decided to build his home in the country. He contacted a business colleague to discuss his plans, and the seeds were sown for a new development outside the village of New Albany. By the late 1980s, these two had founded the New Albany Company and purchased large parcels of land in Plain township with the dream of building a visionary community. Within the next 10 to 20 years, "the greater New Albany area will include some 3,600 to 3,700 new homes, all in one of the numerous American variations
on the Georgian style, at prices ranging from roughly $185,000 to $5 million and up" (Building a dream 1995, p. 60). This visionary community doesn't stop with the planning and building of Georgian homes, however. Careful planning has gone into the schools, the country club, the recreational and cultural amenities, as well as "making sure everything works in harmony with the rural landscape of the area" (p. 59).

As the community grows, so does the school population. The New Albany Company plays an important role in the school system because of its creation of a taxing authority. Knowing that all this development would have an impact on the infrastructure, the developers made an attempt to raise the kind of money necessary to run a quality school district and yet avoid burdening existing taxpayers. In a search for alternatives, the New Albany Company uncovered an instrument which allowed for the creation of a taxing authority permitting the developers to automatically assess a certain millage on land purchased in the development. The New Albany Community Authority was established and a 9.75 mill tax is imposed on any house that is built in their development. "Amortizing that out it comes to about $41 million over 17 to 20 years" (JI1, 4-95, p. 3). People buying new homes pay 9.75 mills in addition to the taxes paid by current residents. Of that money, $28.3 million has been designated for schools, thus allowing for the construction of a high school and middle school campus targeted for completion in 1996. This novel method of
raising money for the schools can and has been a divisive factor between the old and new residents at times. Some, who have lived in this quiet community all their lives, resent these "rich" people coming into their community and making drastic changes. Some of the newcomers feel that since they are paying a larger share in taxes, they should have a greater say in what happens. However, in spite of the disagreements, the voters passed an operating levy and a permanent improvement levy in November, 1994.

Another twist to the story, which has caused considerable friction between residents, has to do with the water and sewer contracts between the city of Columbus and New Albany. The new developments were annexed into the village of New Albany necessitating an increase in the area's sewer and water services. Columbus agreed to supply the service to the annexed areas if it could annex another portion of the township to insure for itself a corridor for growth. In 1994, some residents of Plain township put a merger issue on the ballot which would combine the township and the village. Because of the uncertainty of what impact this would have on the community and people buying in the new development, construction on the new school which began in October, 1993, was halted in early 1994. Both the village of New Albany and the Plain township had to pass the merger in order for it to take place. If the merger went through, however, the water and sewer rates could have increased as much as ten
fold. In November, 1994, the people of the village voted against the merger. The village agreed not to annex a portion of the township. As the agreement stood following negotiations in 1995, the village was allowed to have access to more township land, Columbus could have a portion of the township and the future of another portion of the township was to be determined at a later time. The whole situation caused a lot of tension and resentment among residents of the township. Those feelings were evident throughout the research period among students as well as the teachers who were residents of the community. School construction resumed following the November, 1994 vote and many say the healing has begun. However, as of Spring, 1995, there remained a feeling of betrayal on the part of some residents of the township whose fate was still unsettled, as well as by those living in areas which had been annexed. In addition, some residents perceived attitudes of superiority on the part of many of the newcomers.

School District

In 1991, the school board initiated a search for a new superintendent. According to the current superintendent:

There was frustration among the board; they were getting some rapid turn over, and then in 1992, there was a key board election where the current core was elected. And they had to find a superintendent and find someone who would, 1) try to get some levies passed, and 2) deal with this issue of development and the pressure they were already
receiving because by '92 they had over 230 homes already built out there and these people were building half million dollar homes and they were saying, "I'm paying $8000 a year in taxes," and they'd walk into the chemistry room and there would be only 2 spigots in the chemistry lab. They had no honors classes and the last time somebody passed an AP exam was the early 70s. Their proficiency scores ranked 16th in the county. The Columbus Monthly which ranked high schools came out with a devastating article on what was the worst high school in the county. So the board had instructional types of problems as well as "how do we deal with pressure coming in" and "how do we deal with the old guard." And that was the context in which they were looking for a superintendent. (JII, 4-95, 3)

Dr. Jones, a quiet, serious, and highly motivated high school principal in a nearby, affluent suburb of a large midwestern city was recruited in 1992 as the superintendent of the rural school district to help in the planning and implementation of this visionary learning community. Sherwood, his previous school district, had been involved in the process of restructuring for several years to incorporate new instructional approaches and organizational changes, and had provided many opportunities and inservice sessions for the staff led by nationally known educational consultants on authentic curriculum. Teachers and students at Sherwood had distinguished themselves locally and nationally, and the students had consistently performed well above average on national tests. Team teaching, collaboration, community service, interdisciplinary programs, integration of technology and authentic assessment were part of the teaching repertoire of most of the teachers and administrators at
Sherwood. Dr. Jones, emerging from an innovative, progressive district with personal and professional excitement as well as knowledge of the newest educational approaches, had much to offer the New Albany district. In an interview, he described his role as one of mechanic:

"It's assessment, it's not "got ya." It's modifying, being pragmatic, keeping an eye on the horizon, being a mechanic. ... I have to sort out where my skills are best directed. I've got some responsibility in keeping the board focused and keeping community relations, and finances, so they have the tools to do the wonderful things they aspire to accomplish."

(JI1, 4-95, 9)

The village council, township trustees, school administration, and other community groups all collaborated in collecting ideas for this futuristic community and school system, but the Strategic Plan (Johnson, 1994) was prepared primarily through the efforts of the superintendent, and was approved by the board of education in July, 1993. The plan portrayed the school district's vision of a learning community which was to include a new high school, a senior center, a wellness center, a library branch, a wetlands area, and a daycare center. As explained in a 1993 grant proposal, the learning community:

is intended to provide a setting wherein all community residents and persons working in the community will identify and utilize appropriate resources from the greater metropolitan community to enhance the learning and well-being of everyone, from infants to senior citizens, who reside in the community.... Once the concept of learning community is implemented, there will be a shared sense of responsibility for the education of each child. A learning
community will be characterized by a high degree of community involvement in curriculum design, instruction, and mentoring. The community will serve as the audience for student projects and performances and, in collaboration with the schools, create real opportunities for students to contribute to the community as a whole. The learning community concept will help realize the goal of preparing children for a changing world in which they can realize their full potential. (Venture Capital Application, 1993, p. 1)

The five major goals of the strategic plan are: (1) teaching and learning, (2) assessment, (3) governance, (4) organization, and (5) training and staff development. In addressing the goal of assessment, the strategic plan states that "Assessment must be conducted on which multiple assessments are a major component of a student's transcript and learning process. There is a significant difference between evaluation and assessment. Both must be reconciled. The curriculum is inseparable from assessment" (Johnson, 1994, p. 10). The plan goes on to describe the type of assessment envisioned for the district's students and other points:

• Unauthentic tests are designed to sort students and build in failure. They address primarily the strategy of teacher as wisdom-giver and students as empty vessels to be filled—something someone does to a student instead of with.

• Exhibitions of knowledge should be encouraged, to enable students to demonstrate high standards and competence publicly through a personalized performance that factors out luck as much as possible, such as guessing on multiple choice and true/false tests.
- Assessment is not a terminal occasion for the sorting and ranking of students. Bell-shaped curves that are used for grading and 'curving' ensure mediocrity and reduce collective excellence. Assessment becomes more powerful as an ongoing episode of learning in which students are active participants.

- Shift the definition of assessment from mimicry of the expert's work to intellectual challenges that are engaging, interesting, 'educational' (I learned something from taking the test), and that require the use of judgment and not merely memory or dutiful use of algorithms.

- Technology should be integrated into the learning repertoire.

- Multi-media production capabilities should be used by individual students.

- One's thoughts should be expressed lucidly, orally, electronically, and in writing, for public consumption.

- The limited role of established standardized tests (ACT, SAT, ERB, state proficiency exams, etc.) should be understood. (Johnson, 1994, p. 10)

These goals from the strategic plan illustrate the educational philosophy of the leaders and administrators of the school district.

The three schools comprising the school district stand along a winding road just down from the major intersection of the village. Beside them on a gravel drive, sits a small, gray, converted home which has been expanded to house the district's administrative offices for the school board, superintendent, assistant superintendent and three secretaries. Next door to the
offices stands the red brick elementary building which was constructed in 1956. A driveway and teachers' parking area stand as a buffer between the elementary and middle school buildings. Built in 1925 and the oldest of the three, the two story, brick, middle school is separated from the high school building by the football stadium which sits behind it.

**High School**

The high school, a modular unit popular in the 70s, was erected with the idea that it would be a temporary building; not expected to last more than 15 or 20 years. Flimsy walls, doors, and windows that shake when a plane passes overhead attest to the temporary construction of the building. Separating the rooms are thin partitions which serve as ineffective barriers to sounds next door. Heating and cooling for the building are regulated at a most unlikely location: Atlanta, Georgia. It seems that through a plan to save money by the previous administration, some sort of agreement was struck with a company in Atlanta. This arrangement has proven to be less than satisfactory as fall days turn colder in Ohio, while Atlanta is still enjoying warm or slightly cool days. A visit to a classroom on a frigid, winter day found a teacher and her students huddling around a space heater to keep warm. Windows in the modular unit provide little, if any, insulation against the cold.
Walking through the front door of the high school, one instantly feels constrained, especially if several others are entering at the same time: the trophy cases on the right, the staircase looming straight ahead, and the public phone occupy most of the precious little space. If you venture to the left and into the office, your attention is drawn to a trophy case filled with drill team, wrestling, basketball, and music awards. Covered with more trophies, the counter on the right almost hides the secretaries on the other side. Few can miss Dorothy, the strawberry blond, up-beat, caring and efficient secretary who has a handle on everything that goes on at the high school. From her desk, she has a view of: those entering from the hallway, the principal's office, the counselor's office, and a tiny room, bustling with activity, which houses the teachers' mailboxes and the xerox machine.

Backtracking from the office area to the front entrance, a short staircase ascends to a long hall leading to the cafeteria, teachers' lounge, and the balcony overlooking the gym. The lockers on either side of the hall are covered with signs "Go NA, Beat __," "Good Luck John," and so on. Nearly all the students in the high school are involved in some extracurricular activity in the fall, either football, cheerleading, drill team, or band, leaving the stadium stands almost void of students on a football Friday night.
Entering the teachers' lounge at noontime, one may capture a feeling for a part of the "school culture." Two small rooms make up the lounge area. In the middle of the main room are the lunch table and ten chairs where the high school teachers converge at one of the two lunch periods to eat a quick lunch and spend time with each other. Lively discussions about students, politics, community happenings, classroom problems, or concerns with administrative decisions, as well as personal or even trivial chit-chat among friends, take place in the lounge. Occasionally, the principal joins the teachers for lunch or just sits with them, listening and conversing. The lounge also contains a computer and a phone for the teachers' use. Most of the day, however, the lounge remains empty and emits a cold, uninviting presence.

The band room and a couple other classrooms are located down the corridor past the teachers' lounge and cafeteria. If one turns right after ascending the stairs, there is another hallway leading to the home economics room, the library, a computer lab, and several other classrooms. Located downstairs are the chemistry and biology labs, the art room and business classroom, several other classrooms, and two computer labs.

Individual computers are located in each department chairperson's classroom, the teacher's lounge, and the secretary's office. These computers are networked with the computer labs. Communication can and does take place between teachers and students, as well as administrators, via e-mail.
Two hundred and fifty students attend the high school with a staff of twenty-five. Mr. Kent, a twenty year veteran in education and a very likable person with a charming smile, joined the New Albany staff as high school principal in 1992. He had most recently served as a middle school principal in another nearby suburb. Mr. Kent, views the teachers as very competent and professional. He says he has some 9s and 10s on the innovation scale, but also describes some of the teachers as high resisters. When I first approached Mr. Kent about the prospect of doing research in the school, he was very enthusiastic and supportive. He said that he would like for me to be a role model for the teachers. He wanted them to see that pursuing a Ph.D. was possible and doable and he wanted his teachers to "go for it" if that was their goal.

In an interview, Mr. Kent described his initial impression of the teachers and school at New Albany:

There were a lot of good caring people who were operating as satellites and just sort of out there disconnected. They were doing some good individual things in the classrooms, but there had been a lack of comprehensive leadership to tie them together. It was a typical high school that could have had 1200 kids in it the way they were operating, but they only had 250. (K11, 4-94, 2)

The level of innovation was very low in the school. People were disconnected. People were basically, I think, off in their classrooms, closing their doors, doing their own thing, and if they did something innovative, they didn't tell anyone about it. The cultural change we tried to make is that innovation is OK, it's based on research that is good for kids,
and that reflects good educational practice. The innovation de jour is unacceptable; to innovate just for innovation's sake is not what we are about. What we are about is stuff that is good for kids, reflects good research, and gets good results. (KI1, 4-94, 4)

Mr. Kent's first objective was to join the teachers together in grade level teams to alleviate the disconnectedness. He methodically planned the teacher composition of the teams.

The first time I did teams it was sort of by happy circumstance, but with 10-12 years experience of putting people together, I've learned what works and what doesn't and consequently I've got a high level of innovation in the 9th grade and a high level of innovation in the 12th grade, and I've got some folks in the middle that now are going to have to sort out what are we going to do with 10th grade to help promote that. So it was by design and it was by my experience at trying to put those people together. (KI1, 4-94, 5)

In a middle school and high school staff meeting September 8, 1993, Mr. Kent enthusiastically reported on "What a difference a year makes" and listed twelve accomplishments of the 1992-1993 school year; these included rewriting the master schedule for grades 6 through 12, creating teams of teachers for grades 6 through 12, selecting the team leaders, increasing planning times for the teams, and creating formal and informal communication channels. Changes in the previous year, he indicated, had been primarily managerial changes, but this year (1993-1994) they would concentrate on instructional development concerns.
The change in the master schedule meant that the school day runs from 7:30 until 3:00 with 8 class periods each lasting 50 minutes. Extending the school day by starting at 7:30 instead of 8:00, allowed Mr. Kent to add the team planning time so that each of the grade level teams had a common time to get together each day.

Mr. Kent is enthusiastic about the changes that are being made at New Albany and explains his role in that process. "I think the greatest thing we can do as administrators is make changes and institutionalize them, imbed them in the culture, in the school organization so that it's there. This is the way we do business and it's important for people, for staff members, to become a part of that change" (KI1, 4-94, 6).

When an assessment course was offered for teachers, the principal chose to take it for a couple of reasons.

I wanted to model the behavior that I thought was appropriate for the teachers, which is life-long learning, and learning stuff they needed to know. The second thing was, I wanted to promote a sense of collaboration among the staff; among the members of the class and have a forum for sharing among the levels. (KI1, 4-94, 6)

So even though the facilities are somewhat dated at this point, the ideas and enthusiasm of much of the staff are generally positive and conducive to learning. Part of the "school culture" or values, beliefs, and traditions that this school and the staff
provide are: pride in themselves, and the values of a small community with a personal commitment to education.

During the previous year and continuing into the 1993-94 school year, many changes and innovations from the district level, as well as the school-building level had been proposed and tried. This has led many teachers to feel overwhelmed and overloaded. An example of this occurred on the second day of the 1993-94 school year. Written on the blackboard in the teachers' lounge was a message from the principal:

Staff - You will receive a bus duty schedule (con't from last year) and a NEW teacher absence procedure in your mailbox today. Please place them in your notebook. Thanks::
P.S. Except for the heat, yesterday was great! You all did a great job!

Some of the teachers visiting the lounge at noon were agitated by the message. The teacher who was scheduled for bus duty that afternoon announced that he had gone to the secretary and told her he wasn't sure he could stay after school. Others indicated that this was probably a grievance that should be addressed. One teacher remarked, "We already have been screwed, our school day is longer, how much more do they want?"

Word of the dissension reached the principal's office and he responded at a staff meeting the next week by asking for teachers' concerns. The time element was brought up. The teachers argued that they were already expected to teach classes from 7:30 until 3:00 and that they had no time to prepare for
classes before school. Now they were being told that their after school time was jeopardized, too, with bus duty and teachers' meetings. The time crunch seemed to be a critical issue for them. Mr. Kent reiterated what was said by the superintendent at the staff meeting the day before school started. "We are undergoing a paradigm shift from the factory model and these issues must be discussed and ironed out, not just complained about behind closed doors." He suggested that they needed to think of themselves not as workers punching a time clock, but as professionals doing whatever it takes to get the job done, and asked for a group of teachers to meet with the principals to discuss these concerns.

Some of the teachers were suspicious of all the changes, the replacement of the director of curriculum, and addition of other "outsiders," (a college counselor and new football coach) all of them being hired from the affluent, distinguished district of Sherwood. These teachers sometimes felt that the administrators perceived their school as "not as good as" Sherwood, and that they (the teachers) must therefore be "remade" or changed or even replaced.

The Beginning of Senior Seminar

When Dr. Jones, the new superintendent, arrived in the district, one of his goals was to initiate some kind of senior exit activity. Simultaneously, in the spring of 1993, a high school social studies teacher, Mr. Stern, wondered about creating some
sort of course which would help students with the transition from high school to the "real world." As these two shared their ideas with the principal and others, and more brainstorming occurred, the senior seminar took shape.

Mr. Kent, the high school principal, already had set up teacher teams by grade levels consisting of one teacher from each discipline of math, social studies, science, and English. The senior team was made up of teachers whom Mr. Kent described as:

high innovators, out in front of the pack all the time; that is the way their minds work; that’s the way they are. They were intentionally put on that team to program that team for success even though they are very different people. I got to know them very quickly last year as well as I could and I decided that I wanted my high innovators, my 'thoroughbreds' on that team because they needed to do an excellent job of deciding what the end product was going to be, and that will start driving the curriculum in the earlier grades. So it was by design; it wasn't by happy circumstance. (KI1, 4-94, 4)

Senior Team Teachers

Ms. Dearheart, Mathematics Teacher

Ms. Dearheart has a master's degree in curriculum and instruction and has been teaching for 14 years. A visit to her room reveals much about her as a person as well as a math teacher. Posted on all the walls are her GOALS for the students:

- Understanding the mathematics we do.
- Cooperating and respecting each other.
- Being open to new ideas and new approaches.
• Explaining mathematical ideas clearly.
• Thinking visually.
• Reasoning and thinking about problems.
• Having self-esteem and self-confidence.
• Being able to keep working on a problem.

The week's math assignments are written on the board. On her desk sits a computer terminal which she uses for wordprocessing, for keeping track of students' grades, and to communicate with other teachers, students, and administrators. Ms. Dearheart begins most class periods by going over the previous night's homework, asking for specific problems and having students go to the board to share their answers. During the period, she may show them new concepts on the board or on the overhead projector using her graphing calculator, demonstrating how the concept works in real-life problems. After she poses a problem for the students, she circulates around the room offering suggestions and guidance as necessary. The last 20 minutes of the class are devoted to working on new homework in case the students encounter any problems or questions. The students are allowed and encouraged to ask for help as well as collaborate with others. They must show their work, but are allowed to check their answers to the odd problems with the answers given in the back of the book and the even questions may be checked in the teacher's copy.
Ms. Dearheart describes her role as teacher:

What I see myself as doing is not necessarily teaching math. I teach kids and what I do along the way is that I teach them math. But my prime objective is for them to be self motivated learners and questioning learners. What I want them to do is to be able to operate without someone standing in front of the classroom. (DI1, 9-93, 1)

This petite, energetic teacher works tirelessly to help her students learn, providing the opportunity for retakes on tests so that everyone has the chance to get an A, B or C.

In addition, she is always evaluating her teaching and trying out new ideas and methods:

I never want to become that teacher who is waiting for the day I retire. I can't work in a situation like that. I have to have a challenge. I always have to look at what I can do differently, not even differently; what is it I can choose to do better that will help my students. If I am bored, then I know my students are going to be bored... I try to pick one thing every year that I am going to look at. This year it just happened to be assessment. And because I had seen some things in journals that had kind of caught my eye, and I'm not even sure I'll even get into the area of authentic assessment this year. That may be another year yet before I take a look at that. What I'm looking at basically are alternative forms of assessment, the portfolio, and forcing the kids to look at themselves as the mathematics student and to have them assess where they are and what they think they can do to be a better student. (DI1, 9-93, 3)

For the 1993-94 school year, her "new thing" was having the students do math portfolios. During the summer, she had read
and learned about portfolios, and devised a structured plan for
the students.

In the past, she had her students keep and turn in
notebooks regularly. By checking these, she gained insight into
what the students were learning, but she wasn't sure how much
they learned about themselves in the process. So this year, she
prepared and gave her Algebra II students a packet containing:
the objective of the portfolio, a list of the entries that should go
into the portfolio, the format that should be followed for the
various categories with examples, the evaluation criteria, a list of
her classroom goals, and a copy of the evaluation sheet that would
be used to grade the portfolio. These portfolios would be due at
the end of each nine weeks. So Ms. Dearheart already had some
knowledge of and was using some alternative assessments.

Mr. Stern, Social Studies Teacher

Wearing a warm smile, Mr. Stern, thoughtfully and
enthusiastically confesses, "My dream is to take a bunch of kids
on a bus at the beginning of the year and come back at the end of
the year and I think they would learn quite a bit." When asked,
"What do you think is the purpose of assessment?" he quickly
continues, "That's the kind of thing that changes for me, what
assessment means. I like to judge students in as many ways as I
can. I don't like just one picture of them" (SI1, 1-94, 1).
The slightly receding hairline and first hand knowledge of the Kent State shootings plus his 24 year teaching record at New Albany are the only hints to this caring, energetic, and popular social studies teacher's age. Mr. Stern has a master's degree in social studies plus 25 semester hours. He is currently the chair of the social studies department, the head of the senior team, and serves as the coach for the "In the Know" team, but has served as a student council advisor and baseball coach in the past. He is one of those teachers who can be counted on to tackle a school or community project when all others have said "No." Always looking for novel and interesting ways to involve students in school and community activities, he is the one who first proposed the senior seminar idea which contains a component of community service. Students in Mr. Stern's government class participate in several authentic activities including mock trials, sponsoring and moderating candidates' night, and working on campaigns. These authentic activities aren't something new to Mr. Stern's classes, he has used them for 8 or 9 years.

To help American history students gain a more personal understanding of the Vietnam War, he assigns the students to conduct interviews with their parents or others who lived at that time and find out how they were effected by the war. Mr. Stern utilizes many methods and means to motivate and challenge his students. Films, videos, filmstrips, field trips, questioning, handouts, speakers, games, presentations, projects, group work,
interviews, multiple-choice tests, essays, papers, journals, and discussions are all part of his repertoire of teaching tools, but it isn't just the techniques he employs as much as it is the way he uses them, and the excitement and caring attitude he conveys to his students that makes Mr. Stern an excellent teacher.

Seldom seen at his desk, this dynamo teacher is usually on his feet either in front of the class or walking throughout the room asking questions, offering words of encouragement, or kidding with the students. His classroom has no windows, but the beige colored walls and fluorescent lighting keep it light and cheery. The spinning wheel, weaving loom, antique bicycle, picture of George Washington and two early tombstones offer more detailed glimpses of the subject matter taught here, as well as the teacher.

As the American History students straggle into the second period class, Mr. Stern is handing out sheets of information on the presidents. When the bell rings, he starts, "I want you to look at Ronald Reagan's presidency. What is your impression of it?" Phil volunteers, "He was too old." Mr. Stern continues, "What about the cold war? The build up of the military? The attempt to spend Soviet Union into oblivion, which unfortunately put us into debt or deficit too? Social programs were cut under this conservative era. Now let's evaluate some other presidents together." The students with Mr. Stern's guidance discuss Washington, Jefferson, Lincoln, and Roosevelt. Then he suggests, "Let's list some characteristics or criteria for what constitutes a great president." He picks up the
chalk as one student offers, "Accomplishments, domestic and foreign." Mr. Stern writes on the board as another student adds, "Ability to work with Congress." The students continue to brainstorm: "His appointments," "Their ideals and principles," "Leadership," "Honesty," and "Communication". Mr. Stern then directs the class to the handouts to see what others have thought of certain presidents. "Now I want you to look at Reagan and make up a criteria list of some of the ones we talked about and perhaps some of your own." The bell rings and the students gather their things and proceed to their next classes.

A peek into Mr. Stem's Government class in the fall of 1993 may give some insight into the feelings among the younger residents concerning the changes taking place in their community. As the senior students break up into small groups and begin the process of writing a constitution for the senior class, they complain about the rules and their lack of freedom, and start discussing what they might do about it. One student grumbles about the changes that are occurring in the school because of "all the expensive houses being built and all the rich people moving in and wanting to run things." He is joined by others who offer similar comments. The students are noticeably disturbed.

Ms. Little, English Teacher

Tall, meticulously dressed Ms. Little, her reddish-colored hair carefully arranged in an upswept hairdo, cheerfully greets
the seniors as they file into her fourth period English class. As the bell rings, she hands back their essays and remarks that, "as a whole they were very nice: they were sensible and sensitive." She gracefully moves to the blackboard and raises her neatly manicured hand to write, "You never really finish a piece of writing - you simply relinquish it." She reads her words aloud as she turns to face the class, "If something is underlined on the papers, something is wrong; either a wrong word, a fragment, a run-on. The spell checker won't tell you that b-o-a-r-d is used incorrectly and should be b-o-r-e-d. Look at your papers, count the number of sentences and put that number down somewhere." She picks up a stack of papers written by a syndicated columnist, Stephen Chapman, passes them to the students, and comments, "This is a well done paper. Take a few minutes and read it." After several minutes pass and the students are beginning to look up from the paper, she continues, "What makes this interesting?" One student offers, "The words that make you imagine, like he's wearing a tank top on TV." A lively discussion follows about where casual dressing is acceptable and where it is not. She pulls them back to the topic by asking the question again, "What else makes this interesting?" One student mentions "vocabulary;" another says "humor and examples;" others offer - "people can relate" and "sentence length." Ms. Little comments on their answers and then proceeds with, "Your essays need some variation other than SVO (subject-verb-object format).
your SVO sentences." As she passes out more examples of writings, she says, "Pay particular attention to how to jazz up your sentences. Your job is to revise your papers and meet me in the computer lab on Friday with hand changes, and then revise on the computer to turn them in at the end of the class. Turn in the original and the revised copy. For tomorrow, bring your book and read "The Man in the Water."" The bell rings.

Ms. Little has taught 25 years and been at New Albany for 23. Her jobs include teaching senior English and serving as the school librarian and media specialist. During my research at New Albany, she was completing her practicum and received her Master's degree in 1995.

Mr. Bradford, Science Teacher

Mr. Bradford, the quiet, soft-spoken, science teacher has been at New Albany for ten years. He received his undergraduate degree in criminology, and also has a master's degree in education plus ten hours. Coaching football and teaching chemistry, physics, Biology II and physiology, he sees the purpose of assessment to be a means to evaluate the progress of the students through the curriculum. He uses a variety of methods including personal, one-on-one evaluations, tests, quizzes, oral presentations and video production. For Biology II, Mr. Bradford had the students create a video report on some disease:

I actually had them take a camcorder into the hospital and they videotaped patients who were in treatment or were...
simply convalescing from surgery or what not. That became part of an oral presentation to the class. (B11, 1-94, 3)

An observation in a chemistry class early in the fall found sixteen students sitting four to a table in the chemistry lab and Mr. Bradford walking about the room handing back their graded tests. When the bell rings, he directs the students, "Take the number of multiple-choice you got right and multiply by three and add that to the number on the back for the problems." He then proceeds to go over all the questions giving the correct answers. Mr. Bradford indicates to the students that they will see all of these things again, probably next week, and then moves on to Chapter 6. "The chemical formulas always contain three parts. First, reactants are always on the left side of the equation. The condition necessary for reaction to take place is sometimes included above the arrows. Letters written parenthetically indicate the physical state of the products and reactants," and he shows the various letters used. As Mr. Bradford talks, he writes all of these rules on the blackboard:

Classify chemical reactions.
1. Single replacement reaction
2. Double replacement reaction AX+BY-AY+BX
3. Decomposition reaction AX-A+X
4. Synthesis A+X-AX

He continues "Before you leave today, I want you to balance these equations and I'll check these with your notes," and he writes on the blackboard:
1. \(2\text{Mg} + \text{O}_2 \rightarrow 2\text{MgO}\)
2. \(4\text{Fe} + 3\text{O}_2 \rightarrow 2\text{Fe}_2\text{O}_3\)

As the students begin working on balancing the equations and interacting with one another, discussing the possibilities, Mr. Bradford walks about the room checking their work. A student announces that several of them won't be in class tomorrow as they will be taking the PSAT. Mr. Bradford assures them, "We will review later for the ACT, and the bell rings.

The Senior Team

The teachers who make up the senior team and teach the Senior Seminar are, as the high school principal proudly announced, "his thoroughbreds," "four very good, experienced, academic teachers," "the high innovators" whom he "intentionally put on the team to program that team for success." These teachers possess the qualities of the trailblazers as described by Schlechty (1993). Those leaders who:

are willing to go without maps to places where no person has gone before them, without the benefit of empirically based models, and with little to guide them except belief in themselves, a desire for novelty, the freedom to try, and a vision that motivates and guides them ... Once trailblazers have found a vision in which they believe, all they want and need is encouragement and support for that pursuit. (p. 47)

All of these teachers have continued to take courses over the years, to learn of innovative ideas and practices, and then to incorporate them into their classrooms. According to Mr.
Bradford, "I think anytime you can advance yourself in your field you're better off." They attend workshops and conferences and read educational journals. For instance, Mr. Stern attended a Grant Wiggins' assessment workshop with the superintendent and had seen and experienced first hand how authentic assessment could be used to engage students. Through reading journals and working on the NCTM standards, Ms. Dearheart had developed an interest in math portfolios and authentic activities.

These teachers serve beyond their classrooms; they are all department chairpersons, they teach Advanced Placement courses, they are coaches, advisors, leaders of students' extracurricular activities: football, baseball, In the Know Team, National Honor Society, Odyssey of the Mind, Literary Club, and Student Council. They reflect often on their teaching practices to see where they can do better and they are willing to take risks; to try something new.

Viewed from the outside, they seem to possess the capacities of change agents as described by Fullan (1993): personal vision-building, inquiry, mastery, and collaboration. A personal vision makes their work meaningful. Inquiry encompasses the idea of lifelong learning for the teacher as well as the students. Finally, the ability and willingness to collaborate with others is, as Fullan points out, "becoming one of the core requisites of postmodern society" (p. 14). These teachers work well as a team, capitalizing on each members' strengths, laughing
and kidding with each other, but all focused on creating a meaningful course and experience for their students.

Grades and Ohio and District Requirements for Graduates

In 1990, the State of Ohio instituted minimum competency or proficiency tests to be administered in grades 4, 6, and 9. During the freshman year, students begin taking the Ohio Ninth Grade Proficiency Test which consists of four parts: math; reading; writing; and citizenship. They must continue taking the test until all four parts are passed in order to graduate from high school.

The Twelfth Grade Proficiency Tests went into effect in 1994. Unlike the Ninth Grade Proficiency test which is a minimum competency test, this test is considered a standard's test, and graduation is not contingent upon student performance on this test as it is on the ninth grade test. Students who pass all four parts of this test meet one of the eight criteria necessary for a graduation with honors distinction.

Grades at New Albany are based on an A, B, C, D, and F scale. Students receive report cards every nine weeks. Semester exams are required in courses at the high school level, and constitute one fifth of the total grade for that course. Minimum graduation requirements for the district are passage of all four parts of the Ninth Grade Proficiency Tests, passing of required courses, and satisfactory completion of twenty-one credits.
SECTION TWO
SEQUENCE OF EVENTS

Introduction to Authentic Assessment

When Dr. Jones, the superintendent, created the Strategic Plan (1994) for the district, he described ideas about the kinds of assessment that are more appropriate for the way students learn. He also outlined specific goals and objectives for introducing and implementing those ideas, and how this would be done.

During the summer of 1993, four representatives from New Albany attended a summer institute in New York to introduce the faculty to the use of authentic assessment: Dr. Jones, Mr. Stern, a middle school social studies teachers, and the elementary school principal. Upon returning from the institute, the superintendent met with several representatives from The Ohio State University to discuss the possibility of offering an authentic assessment course for the teachers at New Albany. It was agreed that Dr. Lucas, a professor in measurement at the Ohio State University, would develop a course in alternative approaches to assessment for interested teachers at New Albany. The classes would be held at the high school and a respected liaison from Ohio State would coordinate the arrangements between the district and the university. The superintendent explained the purpose of the course, "to give the faculty the technical skills necessary to do
authentic assessments or alternative assessment, it wasn't just authentic assessments or multiple assessments" (JI1, 4-95, 8).

In the summer of 1993, Dr. Getz was hired as director of curriculum for the district. He had previously taught at Sherwood High School where he had been instrumental in implementing some innovative programs including authentic activities, performances, and assessments. So, in accordance with New Albany's strategic plan and to increase the district's staff with the addition of capable people possessing similar philosophies and knowledge, Dr. Getz was brought on board. According to Dr. Jones, "the first year we changed our director of curriculum, and I began bringing in some key people... That's going to be a five year process, we have time wise when these other key people are going to be brought in" (JI1, 4-95, 6). One of Dr. Getz's jobs was to help in planning the means to introduce authentic assessment to the teachers and to work with them as necessary to implement the activities. I interviewed Dr. Getz and asked what the administration hoped to accomplish through the assessment course. He replied:

If they don't get anything else out of the course, other than a firm understanding of why this is important, we will have made a large step. I hope they understand why being able to do this is important for the students. I hope they understand how this is a mechanism of improving standards ... (and) a mechanism which should help drive some changes that should be occurring in instruction. (SI1, 2-94, 4)
The assessment course was planned to begin in September. At the August staff meeting before the 1993-94 school opening, Dr. Jones talked to the teachers about the changes that were taking place and showed a movie which encouraged risk taking and trying new things. The assessment course was announced and teachers were told to see Dr. Getz if they were interested.

Course Postponed

By September 8th, only 10 teachers out of 60 had submitted fee waivers for the assessment course. A memo was sent to teachers reminding them of the course and asking them to please contact the Ohio State liaison or Dr. Getz if they were interested.

Discussions with several teachers revealed that many of the teachers were feeling overwhelmed with the changes they were working on for this year, and they didn't want to think about anything else right now. By September 17, only 12 teachers had signed up and the decision was made to postpone the course until January.

Staff Meeting setting up the course

On October 18th, 1993, Dr Jones opened the staff meeting by commenting on the groundbreaking for the new high school and middle school, and then continued with a self evaluation of where the school system was and where it was going. He challenged the teachers to think about, "What is good instruction and what do I
have to do?" Many of his remarks reflected his thoughts about assessment in the district and the Senior Seminar course. He reminded the teachers that "assessment is not going to go away;" that "standards need to be open, not hidden;" and so "assessment strikes at the core of what we need to do;" and what needs to be done K through 12, "You can't just have a first class high school; it needs to start lower. You must also have a first class elementary and junior high." He mentioned plans for the preschool and then he directed them to look at the other end, "at a seamless transition from high school to the real world." Next, he alluded to the many areas into which the district needs to put its resources: technology, staff development, math refocus, interdisciplinary teams, vocational prep, multiculturalism, outcomes, graduate work, authentic learning experiences, community service, gifted, Advanced Placement courses, Senior Seminar culminating experience, item analysis of exams, new library, diversity, exit experiences, vocational technology and community involvement. Several possible resources were then suggested: a foundation for raising money, decentralization, and teacher involvement.

At this point in the meeting, Dr. Lucas, the professor from Ohio State was introduced to the staff and continued where Dr. Jones had left off. He emphasized to the teachers that as a school district and as teachers within that district, they needed to know where they have been and where they are going, and must develop a shared consensus of what they are trying to do.
Barriers that would have to be dealt with along the way were time, administration support, education, ideas, parent reactions, resistance by teachers, isolation, expectations and standards, negotiating learning and giving students choices. After discussing demonstrations of learning and exhibitions, he concluded by describing the course he would offer in January. When the first portion of the staff meeting ended, the teachers had been sitting for one and a half hours and were beginning to squirm in their chairs. The elementary teachers returned to their building, and Mr. Kent addressed the middle school and high school teachers. He acknowledged the stress that most of them were feeling due to the new expectations and demands. After this staff meeting, little else was mentioned about the course until it started.

When the course began in January, 34 had signed up: 9 elementary teachers, 9 middle school teachers, 12 high school teachers, the high school guidance counselor, and the principals from all three schools. A questionnaire revealed that sixty-eight per cent of the participating teachers had taught eleven or more years, and that they used a variety of methods to assess their students. The methods mentioned most often by elementary teachers were observations, checklists, projects, short answer, open response, presentations, portfolios and interviews. Middle school teachers used multiple-choice, short answer, true-false, projects, presentations, essays, and observations. Multiple-choice, projects, observation, short answer, essays, and presentations led
the list of high school assessments, but many others were also mentioned: experiments, interviews, performances, portfolios, open response, exhibition, true/false, inventories, matching, simulation, fill-in, oral, and standardized tests.

Many of these teachers already had a basic idea of the meaning of authentic assessment. Their descriptions included terms like "meaningful," "real-life" or "real world," "demonstrating learning," "performance," "process," "audience," "understanding," "other than objective tests," and "natural context." Examples of authentic assessments given by teachers were: (1) "cooperative learning groups taught their sections of the chapter to the rest of the class; students were to present a project by a mode they could choose;" (2) "in general math, students were asked to plan a party for 10 people using a budget, estimating and shopping for it;" (3) "writing a sonnet - students will demonstrate through this creation that they understand iambic pentameter, rhyme scheme, concise use of language, development of ideas; and producing a newsletter." One high school teacher wrote, "All my assessments are authentic."

As the four teachers in the senior team began the course, they had ideas and expectations similar to those of the other teachers in the course, but their understanding of authentic assessment was more concrete. Ms. Dearheart's concept of authentic assessment was "something the kids can really hook onto and say 'well they really do use this stuff outside of class'."
She was already incorporating it into her math classes with portfolios. Mr. Stern had attended the summer workshop where he had seen authentic assessment in action, and had used it in different ways in his classes. Ms. Little's understanding of authentic assessment was that "it is not artificial. That it really demonstrates the skills they have learned and demonstrates them in a way that is real world.... How can they apply it and still demonstrate that they understand it and they have integrated it." Mr. Bradford's conception of authentic assessment was "a realistic, even more valid assessment, than one that is not authentic or less authentic."

The Assessment Course

On January 5th, 1994, the first session of the 2 credit hour assessment course met in the high school library at 3:30. Dr. Lucas arranged the tables in a large U shape with a table for himself and the materials across the open end. When the teachers arrived, they spontaneously arranged themselves around the large U-shape by schools: elementary on the right, middle school in the middle, and high school on the left.

During the first half of the class, Dr. Lucas passed out information cards for the teachers to fill out, a syllabus (Appendix A), and the articles to be read, and then he outlined the three assignments for the course. The first assignment was to create their philosophy of assessment. This had been one of the things
that Dr. Getz had expressed a desire that the teachers derive from the course. The second assignment was to create an individual or team plan of assessment, and the third was interpret, in terms of validity, the standardized scores of three students.

Philosophy of assessment was discussed as Dr. Lucas conveyed information about formal versus informal assessment and the three models of assessment: norm-referenced, criterion-referenced, and the growth model. He also talked about the ways teachers have evaluated students with paper and pencil methods, through products, and by observations of behavior, and the advantages and disadvantages of the various methods. The discussion then turned to self-assessment, meaningful activities, and ownership of learning as a way of leading into creating criteria and standards "with" the students not "for" them. The teachers were invited to brainstorm the criteria which would be used to assess their own assignment, that is, creating a philosophy of assessment. When they reached six criteria and no more were forthcoming, Dr. Lucas had them brainstorm the standards by which the product would be judged. Three articles were assigned to be read for the next class which was to meet in two weeks; the class ended a little past the set time of 6:00.

The next class had to be rescheduled for January 25th due to inclement weather which closed all the schools. I had given Dr. Lucas feedback from the teachers about what they were wanting, so he began by asking the teachers why they thought he had them
write their philosophy of assessment. Someone answered "to articulate your ideas of assessment." Dr. Lucas also offered, "What you are trying to achieve in instruction before thinking of assessment," "how to communicate with parents," and "when I said, 'let's negotiate the criteria and identify the standard,' we were developing a rubric and I was modeling the procedure." He went on to tell them the reason he wanted to negotiate and why he refused to tell them, because he wanted them to own it:

Grades and the kinds of grades one gives impacts on our assessment. It may depend on which model we choose. And then you have to think about what you believe a teacher is. We need to think about what we want students to be able to do as we move into the 21st century. (JL)

The class got into a discussion about where New Albany school was headed. Dr. Lucas spoke of the importance of administrators, teachers, and parents working together and communicating. The teachers then got involved with the discussion and complained about being compared to Sherwood School.

On the Friday following the second class, one teacher commented that if they had another three hours of philosophy Dr. Lucas would have a mutiny on his hands. The teachers expressed their desire to spend time on the kinds of assessment Dr. Lucas had mentioned, and they also wanted to get together with other teachers from different grade levels to share their goals for the students. They indicated they wanted examples, and wanted to be given credit for already knowing a lot of the information they
were getting in the course. They didn't want to be compared with Sherwood; they felt their school could be better than Sherwood. I passed this information on to Dr. Lucas.

At the third assessment course session, Dr. Lucas had the teachers from each group decide who would present their philosophy. They were then to present it to the others as if they were presenting it to a group of parents and the others were to role play parents and decide whether the presenters had met the criteria and standards. Next the teachers were to meet in the three school groups and spend 15 minutes making two lists: the things they wanted the students to know when they got them in their class or what they expected the previous school to have prepared the students to do, and what they would do with the students. The teachers seemed to appreciate the opportunity to deal with the other grade levels.

After break, Dr. Lucas gave examples of different instruction and assessments from classes which he had observed. He offered three examples each for elementary, middle school, and high school. Then he showed a video of two exemplary schools, one in New York and one in Wisconsin. Dr. Lucas also offered some ideas of things that other schools, teachers, and principals were dealing with and some of the problems encountered when using authentic assessments.

Prior to the fourth class session, I mentioned to the teachers that I would like other means of getting feedback from them via
notes to me in a mailbox I had placed in the office and/or e-mail. Dr. Lucas began class by discussing rubrics and how to set them up for a piece of work. He explained holistic and analytic scoring, and said that he was trying to model the procedures for doing this so that the teachers would have ownership of it, just as he hoped they would do for their students. He provided feedback on the philosophies they had created, and talked about taking risks, trying new things, encouraging them to find ways to support themselves and become staff driven. "You need to decide what you want the students to be able to do, then design the instruction, then embed the assessment in that, get away from textbook driven."

He gave the assignment directing them to start thinking about some assessment activity that would last a couple of weeks. and then asked specific teachers from the different schools what they had been considering for that. After several teachers shared ideas and examples, Dr. Lucas showed a transparency of a rubric with a 5 point scale. He described how this is used to help score and to communicate to the students in advance what they are expected to do. He suggested that they use this rubric as a starting point and adapt it to their own tasks. He then continued, "why do we need this? Well if you don't have something when grading essays, you don't get good reliability. For next week, the assignment is to build the instructional activity with instructional intent, method, criteria, standards, and then self assess the
product." A video was then shown which portrayed exemplary schools. It ran past 6:00, and most teachers left before the video ended.

Two e-mail messages were received from teachers on February 22 regarding the course, the first stated:

You asked for comments on the course. For what it's worth:

1) It seems to me that the subject matter must be somewhat limited, since we've heard pretty much the same thing in each session of class. Is there more to this? Or we just not getting it?

2) If the class is supposed to run from 3:30 - 6:00 it's only common courtesy to start and end ON TIME.

3) A little bit of variety in teaching methods (i.e. not all lecture, not all strictly oral) would be nice.

4) the projects assigned are, IMHO, useful and realistic.

5) It's good to have the administrators there.

The next message from one of the teachers who was considered "a resister" was more critical, and the resentment was evident:

Unfortunately I think most of the teachers are very bored with the assessment class. Very little, if any, new information has been given and the lecture or non-stop-soap-boxing gets old quickly. I have been teaching for 26 years and certainly plan a variety of activities for my 45-50 min. classes (let alone for 2 1/2 hour classes that don't end after 2 1/2 hours!). I also try to pre-assess where my students are so I don't insult them by teaching things they already know. All the high school teachers agree that the material covered in class is stuff we've been doing for sometime.

The fifth assessment class met March 2nd. Dr. Lucas had invited 4 students and a teacher from an alternative high school
program who talked about what was unique to their school and how they used authentic activities in the classes. After break, Dr. Lucas discussed indicators and how to communicate indicators to parents and others.

During the last assessment class, March 16, Dr. Lucas told the teachers about the independent study they could sign up to take spring quarter. The independent study entailed implementing their activity and writing up what they found; how it worked, how students reacted, how they participated, what worked, what didn't, and what did the students learn from the experience, and what would they change. Dr. Lucas then had the middle school teachers share the Invention Convention experience, and the senior team talk about the projects and upcoming exhibition. Dr. Lucas pointed out to the senior team that grading might be hard for the judges given the rubrics they had created. He suggested, "Ask the students for help with the indicators and how much things are worth," and noted that in their assignments they seemed to be hesitant in self assessing. "Some of you seemed to be risk takers," mentioning that he saw glimmers of involving parents, community, and school board, and an attempt to get students to present strengths, but hesitancy to let students fail. The rest of the time was spent discussing the standardized test manuals used by the teachers. This was the final class.
"Red-Tape" Problems with the Course

Teachers experienced problems with getting grades for the course. Some got an incomplete for a physical ed course instead of a grade for the assessment course. Three were agitated about this because the teachers get pay adjustments based on hours of coursework towards their masters. This problem had not been adjusted for one of the teachers even a year later.

The teachers at the high school who wanted to sign up for the independent study did not get signed-up on time because they didn't have access to the registration forms in the same way as full time students. They asked if they could do the work spring quarter, but sign up for summer quarter and get the credit then. They were allowed to do this and seven high school teachers elected to take the independent study. Both of these problems had to do with the logistics of conducting a course off campus.

Staff Meeting and Inservice Session

An inservice meeting on February 22, 1994 began with Dr. Jones going over goals and objectives for the district. He tried to bring the staff up to date on the school building project. He talked of standards and benchmarks, and asked them to "consider for the next few weeks and months how do we grow as professionals? How do we get by the conflicts that are inevitable when growing?" He forecast that by 1995, the district would have eight new faculty and by 1997, the faculty would almost double.
In bringing them up to date, he informed them that two weeks earlier a foundation was incorporated and now money would be raised so the faculty could apply for grants (this instrument was created to act as a repository of charitable donations for school improvement). He discussed the concept of a "blur" between school and community which would enable them to tap into businesses, and then mentioned things that are or will be done as a faculty: constantly advancing, Dr. Lucas' course, Dr. Getz getting faculty to appropriate workshops and conferences, item analysis gives us feedback, and then pointed out other plans: technology goals, wellness center, library to be built, wetlands, electronic management to gather data for comparisons with school districts, putting resources into Advanced Placement freshman college courses, new college counselor from Sherwood. The suggestion was made that the senior team talk to middle school teachers to get students prepared for exhibition in earlier grades. Dr. Jones concluded by expressing his extreme pride in what they had done thus far and declaring, "My job is to make your job easier, to run interference."

Dr. Getz then talked to the group about standards and authenticity. He also had them complete a survey on what they felt was important and how much they were implementing at this point.
Senior Seminar

The Senior Seminar course had been created to provide a transitional experience from school to the "real-world" for the seniors. The senior team worked most of the 1993 summer planning activities and discussing what it was they wanted the students to get out of Senior Seminar. The class would meet eighth period everyday so that as the students progressed on their projects they could be excused to do research, or work with a mentor, or whatever else they might need in relation to the projects. The teachers also wanted to do something with the idea of teamwork, so they planned a fieldtrip to Camp Mary Orton, which offers a program to facilitate team building. The students would be required to put together a portfolio and to develop a project throughout the year which would be presented at an exhibition. This was about as far as their plans had advanced when school started in August.

As they began the 1993-94 school year, the senior team felt uneasy due to the fact that there were no specific guidelines on how to proceed. To their knowledge, a course exactly like this was not offered anywhere else. Throughout the year, they felt that they were only a few steps ahead of the students and in some cases not even that.

Since the first year was a pilot and the students would not receive credit for the course, the projects and other activities from the course were worked into the English and Government
semester exam grades. One half of the semester government exam grade for the first and second semesters was based on sixteen hours of community service which had to be documented in a journal. The first semester English exam grade was based on the project proposal plus the student's process journal entries to that point, while the final project, the oral presentation, and the rest of the process journal entries constituted the second semester exam grade.

The students had been told at the beginning of the year that there would be an exhibition in April at which those students producing the projects and presentations judged to be the best by the senior seminar teachers would compete for monetary awards. Prizes donated by the local Businessmen's Association were to be $50 for top presenter in oral presentation, $50 for best product, $50 for best self assessment, $50 for best process journal, $150 for 2nd place overall, $350 for 1st place overall.

The student presentations began April 18, 1994 so that finalists could be selected for the exhibition to be held April 28. The class was divided in two in order to expedite the grading process. Each presentation was scored by three graders. Two of the senior team teachers and the guidance counselor watched and scored half of the student presentations. The other two team teachers along with Ms. Bell, the Jobs for Grad's person who had been hired by the administration to work with the seniors throughout the year, scored the rest of the presentations.
Initially, as part of the assessment course, the teachers had created a rubric for the presentation and set it up as Yes/No (Appendix B). Dr. Lucas' feedback to them consisted of comments on specific parts of the rubric and remarks on the overall rubric. He praised them for their efforts thus far and then challenged them to take it to the next level by suggesting that they:

Start specifying and defining your indicators - use the students to help you through this - they are a good source of ideas and it will be a good learning experience for them. Continue by doing the same thing with standards - get them identified and agreed to.... have the students present more than once - not just the top presenters - by the way, can you clearly articulate to students and parents what will constitute a high level presentation? When students finish these experiences, get their feedback on what worked, what they found helpful, where they had trouble, what they think students in next year's class ought to know before they begin, etc.

The teachers adopted some of his comments and refined the rubric before using it to grade the presentations. They changed from a Yes/No scale to a 0-4 scale. The presentations were scored by all three graders at the time of each presentation, and one member of the team then averaged the scores. Finalists and honorable mentions were then selected by comparing the composite scores.

This first year produced a broad spectrum of projects and quality. Jim built a race car. Marla and Carol gathered information on bulimia, created a brochure, and taped a video in which they told about their personal experiences with the
disorder. Julie created a video which showed her practicing softball and demonstrating the proper techniques of swinging, hitting, and running. Pam organized the class president's notebook. Joan, after working with a photographer for several months learning how to photograph people, created a chronological portfolio of her efforts, pointing out the mistakes in her early work and progressing to more polished, recent efforts. Fred and Tom, working with a landscape architect, designed plans for a memorial garden for the new school, which included doing the actual measuring, planning, and talking to the appropriate officials.

On April 26th, Mr. Stern thanked the seniors for their presentations and announced the finalists who would be presenting Thursday evening at the exhibition. He reported that more money had been donated so there were also seven honorable mentions who would receive $25 each. Folders were passed out for their portfolios.

The 1994 Exhibition

The excitement among the teachers and students was evident as one walked into the high school library April 28th, 1994. The pride of the four teachers showed on their faces and in their words as they talked about the students and made the introductions.
Tension was in the air. It was 6:55 P.M. and hardly anyone was in the small library where the presentations were to be given except the finalists, the judges, and the team teachers. Then, slowly, one by one, they began to trickle in: parents, siblings, the home economics teacher, the school board president, a few juniors, and the seniors who received honorable mention. When the program began at a few minutes after 7:00, most of the 70 seats in the library were filled. "Welcome to the first of what I hope will be many Annual High School Exit Exhibitions" remarked Mr. Kent, the high school principal. He continued:

This began last spring as an idea to have students assimilate their 13 years of learning into some sort of year-long project which would culminate in an exhibition night for the community featuring the 6 best senior projects competing for cash prizes. Four high school teachers comprise the senior team that worked during the summer and this whole school year to make this night possible. I think you will be pleasantly surprised at the results.

The judges were introduced and the exhibition began.

Jim, a shy boy of medium stature took the podium and proceeded to tell how his project took shape and how he used the knowledge he had gained from school. His voice was somewhat monotone, but his grammar was correct and he spoke with no "ah's" and "and's." Jim started from scratch with parts from all over the country and actually assembled a car and built the engine. He hopes to be a race car builder after high school. When it came time for him to show his product to the judges, he led
them outside to where the car was parked beside the school, and
this shy senior instantly took on a different demeanor; his voice
was excited and energetic as he explained what he had done and
proudly displayed the final results. His pride showed as he raised
the hood, caressed the motor and spoke in almost reverent tones
about his car and what all they had been through together. The
judges and boy returned to the library and the other finalists
presented their projects.

Fred and Tom carried their easel and architectural plans to
the front of the room. Fred proceeded to explain how they had
designed a memorial garden for the new, high school grounds. His
presentation was well planned and well articulated. As Fred
described the steps they had taken in their design process, Tom
placed the appropriate drawings on the easel. (Tom's obviously
lesser role in the presentation and overall project had caused
some discussion among the teachers about the grading of team or
group projects.) Their final drawing was quite professional
looking. Fred explained that they had already presented their
idea to the school board and would be trying to raise money for
the project.

After the other finalists' presentations were completed, the
judges retired to another room to deliberate, while the rest of the
audience stretched their legs, enjoyed refreshments, and
socialized. During this time, I spoke with the teachers and they
voiced their excitement about the exhibition and their
disappointment that neither Dr. Jones nor Dr. Getz had attended. Forty-five minutes passed and the judges returned with their decisions and awarded the prizes. Mr. Bradford thanked everyone for coming and expressed the team's pride and enjoyment in working with these students and how much they would miss them.

**Exhibition Grades**

The teachers worked for a week reviewing and conferring with each other in order to provide feedback on the projects to the students, and calculating the grades for the journals, portfolios, projects, and presentations which would constitute their English exam grade. When they finished, the results were presented in class. They gave each student a packet which included a sheet showing who graded them and what the total score was, the individual evaluation sheets of each scorer, and finally, a sheet of general comments prepared by the teachers.

**First Year in Retrospect**

This was the first year, a pilot. The students had no exemplars, no concrete examples of what was expected. Those students who received low grades were upset and disappointed. They felt that the expectations had not been clear and that they had done the kind of projects that had been acceptable throughout their school years and all of a sudden it wasn't good enough.
Reflecting on the year, the teachers felt the process journal should be standardized for the following year. The student skills that needed work were research, presentational, and video production, and so they made the suggestion that research and oral speaking skills be written into the 9th grade curriculum. In the senior team's May, 1994 interview, the teachers commented about the assessments they had used:

Mr. Stern: We were hoping that all of this would be authentic. Now some parts of it are authentic and some parts of it are as phony as all get out.

Ms. Dearheart: I think we need to sit down with the kids and have them help us develop the rubrics. Most of the rubrics we kind of bounced things off them [the students], but I think we need to get a little more specific than that and ask the kids to develop what it is they want us to look at.

Ms. Little: Our rubrics were too generalized, yet it was better than nothing. The rubric we used for evaluating this year will have to be changed because it was repetitive.

Mr. Stern: Next year... I think we're going to have a clear picture in mind of what we are looking for because we had no clue. We had no clue going in as to what we were going to get. We really didn't know how these presentations were going to look. So next year we will have videotape. We know what they were like this year. We will be able to say, "Well this is a dynamite presentation; here's some things that you should be looking for." I think we can give some do's and don'ts.

Ms. Dearheart: With all of the pieces except community service - that's going to be at the end - but the other two pieces of the seminar (team building and careers) are going
to feed directly into the projects. We're going to expect some understanding. (TI1, 5-94)

They also discussed the need to change the percentages that were attributed to each activity:

Ms. Little: One thing, I was not happy with the process journal being only 20% of the second nine weeks.

Ms. Dearheart: Yes I think in retrospect we should have made that more.

Ms. Little: Well, you know when we set those percentages there, it didn't occur to me that 20% is fairly inconsequential.

Mr. Stern: And some kids' products were so much better than others that you were comparing apples and oranges. (TI1, 5-94)

Once the projects with the seniors were all completed, the teachers edited a videotape of the first year's presentations and prepared a talk about Senior Seminar that they gave to the juniors in May. They talked about this year's program and their expectations for next year. They showed excerpts of the video tape to the students.

During the first year of Senior Seminar, the teachers worked very closely with the seniors throughout the year helping them think about what they had learned over the years, what they would be doing in the future, and how to integrate that into a project. A bond developed between the teachers and the seniors. The students signed a petition for the senior team teachers to
hand out the diplomas at graduation, but since that was not deemed appropriate, the teachers were allowed to read the students' names and hug them as they came off the stage. Mr. Kent described the intense emotion of the 1994 graduation to the juniors the first day of school in August:

The emotion between teachers and students at graduation was incredible. It was unlike anything I've ever seen in my career. I think you all can develop this same relationship.

June, 1994 Inservice

At the end of the 1993-1994 school year, the administrators provided a three day inservice session for the teachers. The theme was "If we spend time in planning, the process goes better." They wanted to create an atmosphere where the teachers could reflect and plan. Dr. Getz talked to the teachers about assessment and the proficiency tests, of making instruction and assessment more authentic, about how each building had created its own philosophy, about bringing in parents for community involvement, and the concept of a learning community. Acknowledgment was given to those teachers who were taking risks and trying out new ideas when he reviewed all the various projects, workshops, and programs that have or were being undertaken within the district.

The senior team used the inservice time to critique and reflect on the year's program and then to plan for next year by developing the graded course of study. They continued to work a couple more days after the inservice, and then they each had
individual assignments to complete before a July deadline. They presented the Senior Seminar Graded Course of Study to the board in July and it was voted on as a graded course of study as well as a graduation requirement at the August board meeting.

SECOND YEAR 1994-95

At the staff meeting the day before the 1994-95 school year, August 29, Dr. Jones explained how authentic assessment can clarify our goals. Dr. Getz then discussed staff development and challenged the teachers to identify areas in which they needed personal development and to work these into the professional development planning guide for implementation of the Venture Capital grants.

Over the summer, the Senior Seminar team teachers did a lot of reflecting on the planning, implementation, and outcomes of the previous year. Based on these reflections and discussions, they made substantial changes and set new expectations for themselves and the students for the second year.

The first day of school, the teachers spent first period, which had been designated as the team's planning time for the 1994-1995 school year, and eighth period, which is the regular senior seminar period, welcoming the students, explaining the goals of the course, and their expectations of the students. Since the
course had been changed considerably and was now a requirement, the teachers began by telling the students just that:

We promised that we would get this course for credit and we did this summer. So whatever you heard about last year, this will be completely different. This course is very important to us and we hope it will be for you, too. It will be well organized as an exit learning experience. (O, 8-94)

The principal was present at the first class and challenged the students, "In the past 12 years, we have done things to you, now we want you to participate in your own learning. That is why this class is eighth period, if you need to be dismissed early for mentoring or other educational opportunities" (O. JL).

The teachers passed out the student handbooks and went over some of the rules. Notebooks for keeping all their materials for Senior Seminar were given to all the seniors. Everything that the students received in the way of papers or handouts were color coded: white-general information, yellow-career nexus, blue-team building, orange-project, white-community service. To standardize the assessment and control the length of the journal entries, the teachers also purchased journals for the students. Ms. Little explained to the students, "This makes it a little more uniform than last year and when we say write a page, you know what we mean."

The teachers had also decided to implement the concept of "redos" which would enable each student to get at least a B on every assignment, but required each student to redo and resubmit
work until it met that standard. "Our intention is that no one gets below a B. If we decide it doesn't meet the standard we have told you, we will give it back to redo. This will be work "in progress," if not completed after 9 weeks, you will get a zero." Each teacher had selected the quarter for which he/she would be primarily responsible. They took turns explaining their particular focus, what would be expected, and how it would be graded.

During the second week of school, the teachers conducted a Senior Seminar parent meeting to describe the course and expectations to parents. Pamphlets created by the teachers explained the philosophy and the activities related to the course. The philosophy statement, included in the pamphlet, had been the first assignment for Dr. Lucas' course. It had been refined for the second assignment, and was slightly modified for the second year's Senior Seminar course (Appendix C).

Mr. Stern explained to the parents the purpose of Senior Seminar and how it was conceived. He emphasized that "we are very careful about how we are assessing the students. We want it to be very clear to them, so there is not confusion." He explained how the grades would be derived for the course and discussed the projects and exhibition. The meeting seemed to go very well and there was no negative feedback from parents.

Second Year Exhibition, April 1995

As I entered the library, the words that echoed in my head were, What a difference a year makes! Student projects spilled
out into the hallway. Students were busily putting last minute finishes on their displays, tables were set with catered pastries, a video projector, a computer, and LCD panel were set up front for student presentations, an impressive, white podium stood at the front of the room, and tables were arranged in a semi circle around the podium with chairs for the nine judges. The programs which were handed out at the door by a student listed the judges who represented various positions in the community: a Pastor, the district superintendent, three members of the New Albany Businessmen's Association, the Mayor, the Postmaster, a representative from the New Albany Company, and a village council member. Further reading of the program revealed that there were six finalists and fourteen honorable mentions this year.

The program began with a few remarks by Mr. Kent and then the exhibition got underway. The presentations by the finalists consisted of the restoration of a '55 Chevy, a jazz arrangement, a room construction, a golf course design, a computer program, a music ministry promotional package, and a recording and label design for a rock band. Examples of projects done by those receiving an honorable mention were an antique show arranged and conducted by a student, a stage adaptation of "Beauty and the Beast," a Senior Video Yearbook, and an expressway impact study.
Following the presentations, the judges marked their scoring rubrics which had been given to them a few days in advance. During the break, the scores were tallied on the computer by Ms. Dearheart. Meanwhile, the presenters and the audience enjoyed refreshments, looked at the other projects, and socialized. Within a half an hour, the results were announced and the prizes awarded. The comments made about this year's exhibition and projects were very positive. One judge remarked that he wished he still had children in school so that they might benefit from this kind of experience. The teachers were also pleased as illustrated in an interview a few weeks after the second exhibition:

Ms. Dearheart: In general, the projects were better, they were better because we knew more what we wanted and we were able to communicate that to the kids.

Mr. Stern: I think we did more preparation for the kids this year. More than we did before and helped them a little bit more putting things together; understanding how to do that.

Ms. Little: They also had time to think about it, though, because they knew they were going to have to do this. So a lot of them had given thought to it over the summer or started over the summer or whatever. We didn't have so many blank looks when we conferred with them individually anyway. At least they had thought about it.

Mr. Stern: And the students did a much better job with presentation skills this year.

Mr. Bradford: Technically much more polished: Video editing, audio editing.
Mr. Stern: Some of them were embarrassing last year, some of the presentations. The fact that they weren't finished

Ms. Dearheart: We had a couple this year.

Mr. Stern: You're always going to have that. I think most of our honorable mentions this year would probably have been winners last year.

Researcher: To get the polished videos and things, didn't you work on skills like that throughout the year?

Mr. Stern: A lot more

Researcher: And presentational skills?

Mr. Stern: I think the structure helped us too to do that. There was more. The structure allowed us to spend more time with them that way. We're even thinking how we are going to do that. (TI3, 5-95)

Overall, the senior team teachers were pleased with the course and the assessments they had developed. They reflected on this year and planned for next year during the June inservice, and then requested an additional two days beyond that.
CHAPTER 5
FINDINGS AND DISCUSSION

In this chapter, I will address the questions which guided this study and describe the problems, issues, and concerns that were encountered as the teachers were introduced to and tried to implement authentic assessments. These include time, support, communication, reliability, the overall goal of the district, and the teachers' place in that plan. These issues will be described from the perspectives of the teachers, as well as the administrators and students, when relevant, and then any discrepancies will be discussed and analyzed.

The theoretical bases for this study are change and constructivism. These are apparent at several levels throughout the study: at the classroom level with student learning, at the level of teachers changing approaches, at the level of teacher-principal interactions, and at higher administrative levels. Accordingly, a major theme which emerged during the research was the different viewpoints or perceptions of the participants occupying various roles and positions within the school, that is, the teachers, the administrators, and the students. Many of the situations that I encountered took on a different focus based on
the role of the person with whom I spoke. Not only did the viewpoints differ from position to position in a given situation, but many times within the positions or roles as well. Therefore, the research questions will be addressed at several levels. First, I will try to answer them from each of the viewpoints relevant to that question. Next, I will discuss the different viewpoints and will compare what is happening in this research to what previous research and related literature has shown.

**QUESTION 1: How does a school district manage the introduction and implementation of authentic assessment into the classrooms?**

According to Fullan (with Stiegelbauer 1991), "The district administrator is the single most important individual for setting the expectation and tone of the pattern of change within the local district" (p. 191). This seems to have been the case at New Albany as the teachers were formally introduced to authentic assessment through administrative decisions and specific plans. Upon being hired as superintendent, Dr. Jones was instructed by the school board to create a plan for the district. Though well aware of the need for teacher and community input for implementing change, Dr. Jones explained the reason for the district's top down approach:

There is an interesting story behind that. You've heard of strategic planning. And in [Sherwood], we plan, collaborate, cooperate. So we got to the point here where I came on
board, and the board said, "This is what we want, now how do we get there?" So I came back and said, "You need a plan and you need a roadmap on how you are going to do it. We'll get some community meetings together and some faculty and we'll do a strategic plan process and make sure everyone is involved." And they looked at me like I was from Mars. They said, "Well we don't want to go to any meetings. If you get people together right now, there is going to be friction. You know where the faculty is at. We hired you, you bring us back a plan in another month and if we like it, we'll adopt it and then you get to work on the 31st day." ... So I went away for a month, put my beliefs down on paper, shared it with the faculty, and said, "These are the things I got and I'm going to the board with it, and the train's leaving the station." (JI1, 4-95)

The resulting Strategic Plan (Johnson, 1994) which defines the district's goals and objectives for carrying out the plan fills a very large notebook in Dr. Jones' office, and includes work plans for the principals, the director of curriculum, and the superintendent. These people meet weekly and report on their progress. The formal introduction to the staff was described by Dr. Jones in an interview, "It was primarily through meetings, entire district meetings that I held with the faculty, outlining what some of these [the goals, objectives, and work plans] are." The introduction of authentic assessment into the district was one aspect of the overall plan. Mr. Kent, the high school principal, recalled those meetings with Dr. Jones shortly after he was hired as superintendent, discussing alternative assessment strategies, and then gathering and reviewing all of the high school semester exams. He commented that what he found was:
a dismal array of standardized tests taken directly from masters of textbooks, to teacher-made tests that were hand written, to teacher-made tests that were just not real cohesive in their approach. All of the examples that we had were typical paper and pencil kinds of things for the entire testing period.... I collected them and sent them over to Dr. Jones and he sent them all back to me with a note that said, "we've got a lot of work to do." (KI, 5-95)

Mr. Kent created a rubric for the exams, scored them, and returned them to the teachers with feedback. This process has continued for the past two years.

Stiggins (1995) has proposed an assessment literacy plan for school districts. He agrees with Fullan that a move to assessment literacy begins in the highest positions of authority. He defines the first step as the articulation of clear and specific achievement expectations of students, teachers, and the district. New Albany's district's plans were well laid out in the Strategic Plan (Johnson, 1994), but in terms of student expectations, the vision of what the successful graduate needs to know and be able to do was initially somewhat hazy. That vision, however, is being clarified through the efforts of the Senior Seminar, and the expectations for the students are beginning to be established and shared with teachers throughout the district. So the district's efforts seem to be on target in accordance with the first part of Stiggins' plan. However, in terms of the literature which suggests involving the teachers in the planning, the district seems to have initially omitted this step due to the differing perceptions of the board and superintendent.
So the changes to the district were top down beginning with the school board.

**Workshop**

The next formal step taken by the district to introduce authentic assessment to the teachers was to "plant seeds" and have some "key people" (JI,4-95) attend workshops. During the summer of 1993, Dr. Jones arranged for himself, two teachers and the middle school principal to attend a workshop by Grant Wiggins in Geneseo, New York. Dr. Jones' previous school had been using authentic assessments for several years and had worked with Wiggins on several occasions. According to Dr. Jones:

> You know when you hear Grant Wiggins, you get a whole different perspective in terms of how to do assessment. I planted seeds among the faculty. You want someone in the audience nodding their head saying "yeah, we need to be looking at these types of things if we are to grow professionally." (JI1, p. 8, 4-95)

Mr. Stern, the high school social studies teacher who attended, explained that they were sent to the workshop to generate the spark and put someone in the position of being enthusiastic and supporting the idea of authentic assessment when it was brought up in various situations. Mr. Stern had previously used authentic activities in his teaching and returned from the workshop excited and enthusiastic about what he had seen. He was in a good position, chair of the social studies department and head of the senior team, to have input and share
ideas about what he had learned concerning authentic assessment. The other teacher who was sent to the workshop was a middle-school social studies teacher for 8th and 9th grade. The principal who attended the workshop had been the elementary school principal, but took over the position as middle school principal for the 1993-94 school year.

Was this workshop approach consistent with what the literature suggests? In addition to involving the teachers in all aspects of the planning of changes or innovations, the literature on change and staff development also suggests involving teachers and staff members in the introduction and implementation stages. Dr. Jones' inclusion of a principal and a couple of teachers in the first step to build a common understanding and interest was in line with what the literature recommends. However, because the principal was now at the middle school, there was no representative from the elementary school who could carry the ideas and enthusiasm back to that school. The number of representatives was very small, but the teachers were as Dr. Jones noted "key people" in leadership roles who worked well with the other teachers and would share their ideas and spread their zeal.

It was a one shot workshop away from the everyday routine of the school in contrast to what the experts recommend, but it was meant to be only a beginning, not the sole exposure to authentic assessment. It served the superintendent's purpose of
familiarizing and enticing several of the staff members with the ideas about authentic assessment prior to the assessment course.

**Assessment Course**

The next major step was the arrangement of the assessment course. However, the administrators' expectations of the course for the teachers were different from the teachers' expectations, and this may have contributed to some disillusionment and resentment among the teacher participants as well as resistance.

**Administrators**

The superintendent explained that the purpose of the course was to "give the faculty ... the technical skills necessary to do authentic assessments or alternative assessment" (JI1, 4-95) and to help the teachers develop an understanding of assessment. As mentioned in Chapter 4, Dr. Getz noted that he hoped the teachers would get an understanding of why this was important to the students as well as an understanding of how this can improve standards. He added:

I hope they have some good examples in terms of how, but I don't think that there are any..... I hope that they get the realization that this can make teaching much less of a mechanistic kind of thing; that it is more of an intellectually engaged activity that they are constantly evaluating the teaching-learning process and that this kind of assessment is a goal...... They have to be able to allow themselves to take the risk of experimenting and learning along with the students as they do this.... So if they don't get anything else
out of the course, other than a firm understanding of why this is important, we will have made a large step. (GI1, 2-94)

The high school principal viewed the purpose of the course similarly to the central office administrators. Mr. Kent explained, "My desire was for people to get a different set of glasses to bring things more clearly into focus" (KI1, 4-94).

Teachers

The teachers who signed up for the course listed several expectations they had of the course in a questionnaire. They mentioned that they wanted to learn different and better methods of assessing students effectively, to learn more about authentic assessment, to see real examples of authentic assessment, and to learn how to discuss students' progress with parents (For a complete listing of teachers' expectations, see Appendix D).

The senior team's expectations were similar to the other teachers. In individual interviews, the teachers addressed specifically what they hoped to gain from the course. Mr. Stern asked, "What can I do with paper and pencil that is authentic, if there is such a thing?" and indicated that, "What I need to do and what this course might help me with is to develop more careful rubrics to judge things with" (SI1, 1-94).

Ms. Little hoped "To see if there is anything else I can do to better hone my skills at assessing. I hope he [Dr. Lucas] will give us some new methods or modeling. Let's see you do something or
see a tape of somebody who is doing it." She also indicated that she wanted to "learn more about self assessment and how to create rubrics, and some tried and true type things that somebody... says, that through research, these are the things that are working, and when to use each kind of assessment and how many different kinds to use" (LI1, 1-94).

Mr. Bradford had several reasons for taking the course. First of all, he felt that taking courses at the high school with the other teachers "develops more camaraderie with other people within the district." Another reason was "you never know when you are going to pick up a good idea. I think any time you can advance yourself in your field you're better off." The final reason he offered was that it was free. If he had to pay for it himself, he might evaluate it more carefully. Through the course, he hoped to learn what exactly was meant by authentic assessment:

what new ideas there are and to see how different grades below mine are evaluating the kids so I know what to expect from the kids when they get to me. (BI1, 1-94)

He indicated that he might be able to pick up some things that have worked for others that he might try, "perhaps certain techniques of test structure or handling grading, that is to say, with video presentations." (BI1, 1-94).

The expectations of the teachers taking the course were generally of the practical, real-world type. They wanted examples of assessments they could use in their classrooms. They felt that
they already understood the need for alternatives and were ready to learn what those alternatives were. After the first couple of weeks, the teachers complained that the administrators and the professor of the course did not give them enough credit for what they already knew, since they saw much of the course time being spent on things like philosophy. They complained that they already understood these things or that they wanted examples and modeling of the ideas, and they wanted to know how to go about creating rubrics.

The teachers' expectations were in conflict with those of the administrators who believed the teachers needed an overall change in beliefs and philosophy about learning, instruction, and assessment. Most of the things the teachers wanted were shared, demonstrated, and modeled throughout the course, but for some reason that is not what the teachers perceived. The professor modeled how to create the rubrics by working with the teachers to create the criteria and rubrics for the assessment class assignments. He provided examples from several schools, in his lectures, in videos shown to the class, and by having a teacher and students visit from a school which utilized such alternative assessments in their program. The professor also tried to meet the needs and wants of the teachers by responding to feedback passed on to him by me, the researcher, such as providing examples and giving the teachers time to share with each other.
Mr. Kent, the principal, sensed that the teachers' objectives were different than the administrators in an interview following the course:

We needed to have more authentic examples of authentic assessment, and I think that was the desire of most of the people, many of the people in the class, even though we had some really wonderful examples articulated by different levels. I think the staff was feeling they wanted a silver bullet; they wanted to know the formula and like so many in our business it's a way of viewing the world. (KI1, 4-94)

He attempted to explain that the difference may result from the teachers not understanding the change process:

I think, people, generally, on the staff don't understand the change process, don't understand what it takes to make a cultural change in the way we do business. And like I said they wanted a cookbook approach to authentic assessment, you do this first, then you do this, and then you do this, and here's your result. And it just doesn't happen that way, they don't understand yet, many of them don't understand that professional judgment, inquiry, reflective thinking, documentation of a case and following through and thinking about it and then adjusting your behavior after you've had an opportunity to think through all that and reflect on that is the thoughtful, professional growth that needs to happen. And unfortunately, most people, many people on this staff are at ... the basic control and safety level as opposed to the level of self actualization. And I think that is totally predictable because a lot of these folks have been in high school education for a long time and they have been beaten down by folks who just didn't want them to think about stuff, they didn't want them to reflect on that stuff. (KI1, 4-94)
Did the assessment course accomplish what the various groups had hoped?

From the district administrator's perspective

I feel that they (the teachers) are more comfortable with the "why" than what they were. They have a good sense of that. The next step is the "how." Until they actually start doing that, you won't see the instructional shift as much ... while on the rating of perceived importance the teachers are rating that as high, in terms of implementation, they are rating it as low. So we're not having that shift. (GI2, 8-94)

This statement was made at the beginning of the 1994-95 school year in an interview with Dr. Getz. He mentioned that as he had talked with the teachers in the spring and they reflected on the strategic plan together, they "either explicitly stated or implied some frustration about having models and exemplars." The course probably accomplished the basic objectives of the course as described by the administration, that is, helping the teachers to develop a philosophy of assessment and to understand why doing this is important, and to give them the basic technical skills to do authentic assessments. (However, it needs to be kept in mind that this course met only fifteen hours, thus the depth of the course in terms of technical skills was limited.)

From the principal's perspective

"It [the class] was the formal opportunity to learn the talk and all that stuff, it was not the class by itself that had an impact, the
class was part of the mix, the eclectic approach." (KI2, 4-95) He also acknowledged that the teachers had expected something different; they wanted, as he described it, "a silver bullet."

From the teachers' perspective

The questionnaire responses of the high school teachers revealed a variety of answers to the question, Did you learn what you hoped you would learn about assessment?

No. Everything that was presented, we had had presented earlier. The concepts that were presented could have been presented in one or two classes.

No.

No. But what I DID learn was how to develop better standards to evaluate students. That's not what I envisioned, but it certainly helped!

Yes. Mainly, the course gave me some tools for making up my own methods of assessing.

No. Nothing was really "new" information, although I did enjoy hearing about actual student's projects and exhibitions.

No. There weren't any new ways of assessing discussed except for the self-assessing.

Not completely. We belabored the obvious too long- hoped to see or hear some state of the art programs in Ohio or U.S.

Yes. I believe that the conclusions I had come to in my own teaching experiences regarding "tests" were validated.

No. Most of what was offered I've already had many times in other courses working on M.A.

From the senior team teachers' perspective

Since the senior team teachers came to the course with some previous knowledge and experience with authentic assessment
and the course was gauged for all teachers in the district who chose to take it, what did the team gain from it? In a team interview they explained:

Mr. Stern: It probably didn't help us in terms of giving us a lot more information, although we did know some of those things and some of it we didn't. I think what it did was jump start us. It gave us a reason to sit down and start planning. (TI1, 5-94, p. 1)

Ms. Little: This was not exactly a big eureka thing to us ... We were already doing that. The only thing I learned was the self-assessment thing. I had never really given the kids an opportunity, but I did that and now what I've learned from that is ... you have to lead them through that. But I don't know how to lead them, how much prompting to give them until it becomes my answer and not theirs. (TI1, 5-94, 6)

Mr. Stern: The level of teaching we're doing now and what we've heard already, you can't do a general course because it's absolutely useless; all that course is intended to do is to give you an overview and motivate you, which we're already motivated and the overview, we have a wealth of our own independently. What we needed to do is what we are doing right here, sit together and work on stuff, and I think if the course was designed so that we would have specific tasks to do that were related to better assessment and our groups, I think we would have gotten more out of it. (TI1, 5-94, 10)

Ms. Little: He didn't deal very much at all with developing a rubric just kind of said, "Well that's your assignment and then go do it" and stuff, that's kind of new to us. (TI1, 5-94, 11)

Ms. Dearheart: We needed more guidance. How do you lead students through? (TI1, 5-94, 11)
Mr. Bradford: I thought we were going to see examples of successful programs using authentic assessment ... even to the point of traveling somewhere to see something. (T II, 5-94, 11)

In the senior team's view, the course was too basic and general for them, but it did give them a reason to sit down and start planning, thinking through their ideas and philosophy, and then to prepare a statement of what the Senior Seminar was designed to do and how they planned to accomplish those goals. Although the senior team teachers initially complained about having to think about and write out their philosophy, when they planned for the second year of Senior Seminar, they were glad to have their ideas on paper to share with others (for example, parents, students, or school board members). The course and the professor’s feedback also provided some guidance in creating the scoring rubrics for the assessments in the course. This was not necessarily all that occurred, but the teachers' perception.

Other Formal Introductions to Authentic Assessment

Authentic assessment was also introduced to the teachers at staff meetings and during an inservice in which Dr. Getz talked about standards and a continuum of authenticity that he had developed. The middle school principal and the two teachers who attended the Wiggins' workshop were in these meetings and the inservice to testify to the benefits of authentic assessment that they had seen in action. The senior team teachers commented
that they had learned a lot about assessment from Dr. Getz, and Mrs. Dearheart added, "I remember thinking the same thing and thinking 'Gosh we should hear more from Dr. Getz, he has a lot to offer us and a lot to tell us." Dr. Getz also purchased several books on assessment for the staff.

What about the resistance and tension among the teachers? Were things done during the change process that contributed to that? What could have been done differently to avoid the tension and resentment?

Mr. Kent indicated that some teachers were resistant to the changes:

I've got some high resisters who basically, for a variety of reasons, it's either their personal life and what they've got going on in their personal life, or they're three years from retirement and they really don't feel like changing or they really don't buy into it. They don't believe in this authentic assessment stuff. And it has absolutely nothing to do with gender, time in service, anything you typically would expect. I've got beginning teachers that are as, or more rigid, than folks who have ten years experience; I've got people with twenty-five years of experience that are high innovators and I've got people with twenty-five years experience with absolutely no innovation. So it has to do with philosophy and belief and what drives you, as opposed to years experience, gender, race, sex, any of those things. (K1, 4-94, p. 5)

There were resisters within the district, but most of teachers taking the assessment course including those considered the innovators had similar perceptions of the overall changes within
the district. There were many changes being brought about at once: newcomers into the community, the problems and tension between the village and the township and Columbus, a new superintendent, a new director of curriculum, a new college counselor, talk of new teachers coming in (all of these from Sherwood), the principal collecting and reviewing the semester exams, the item analysis the teachers were expected to do on the exams, the push for more Advanced Placement classes, the proficiency tests, and more. The lack of involvement in the changes may have contributed to the teachers' perceptions that the administrators did not think they knew anything, their resentment, their frustration, and sense of being overwhelmed, and their perception that they were being compared to the teachers at Sherwood and feeling threatened by that. Although the assessment course was neither mandated nor required, some teachers felt the need to convey their interest and enthusiasm to the new administration.

Central office was aware of the tension and resistance among some of the teachers. In an interview with Dr. Getz towards the end of the 1994-1995 school year, he acknowledged the tension and felt that it had been dispelled somewhat, but noted:

Anytime you're in a change process and you see its going to be continual ... there is always going to be a certain level of tension or anxiety and there almost has to be a certain level of tension to get you to change in the first place.... There is some, it's still at a reasonable level; we'd like to say we have
everybody on board, we don't. There are still people who
are resisters or there are still people who will talk the right
language and so forth, but really for reasons that are varied,
not moving in that direction in their own personal growth or
in their classroom instruction, .... What I think that I see
emerging is a very, very strong climate of wanting to discuss
these things. It's almost like, it's now more the in thing to
be a participant than a resister. (GI3, 4-95)

The administrators, both at the district level and the high
school building level, have worked to bring the resisters on board.
They have been supportive and encouraging of those who were
the trailblazers, their "thoroughbreds," which at times caused hard
feelings and "a large jealously factor," but the administrators soon
recognized these problems and have tried to find areas in which
the resisters could also find their niche, become involved and
form more cohesive teams like the senior team. By the spring of
1995, even the two most vocal resisters had become involved
with a team project and were quite enthusiastic about it.

District's Future Plans for Authentic Assessment

The Strategic Plan for New Albany in terms of authentic
assessment should be completed by the 1996-1997 school year.
According to Dr. Jones:

We wrapped it into the teacher evaluation component too,
so that where part of the strategic plan indicates that by 96-
97 every teacher shall be, (one) trained in authentic
assessment, and there is a work plan on that, and (second
phase) is, it is incorporated into actual classroom activities
which should be observable by the principal coming into the
classroom. By 96, not only by observation, but by item analysis of the teachers midterms and finals. (JI1, 4-95)

Although the central office administrators' initial ideas were based on their perceptions of what the teachers needed, and they may not even have been aware of the teachers' resentments at that time, their follow up plans for professional development and the implementation of authentic assessments are based on the teachers' needs and wants and having the teachers continue to think about "why" they are doing this. In an interview before the 1994-95 school year started, Dr. Getz explained that the teachers were not implementing authentic assessment much yet, and that they wanted models and exemplars to follow:

I think that even though a lot of them now are buying into the "why," and they have done some of this kind of activity; the activity is still not pervasive enough to cause a sustained shift in the instructional approach. The other thing that relates to standards... is that we find that there are two or three real sources of standards. One of the sources is your curriculum, your graded course of study.... So I want to bring that into alignment... but an additional component is models and exemplars and as some of the teachers reviewed this strategic plan, and reflected on the activities we did last spring, there were a number of things that either explicitly stated or implied some frustration about having those models. So that is going to become a new strategic plan initiative. (GI2, 8-94)

The results of the Implementation Survey given to the teachers during the year by Dr. Getz had indicated that with the questions having to do with assessment, teachers rated importance from 4.6
to 5 on a 5 point scale. But their rating of the extent to which their building presently reflects the attributes ranged from 2.8 to 3.9, clearly indicating room for improvement. Dr. Getz explained what needs to occur:

To work on the how of authentic assessment and continue to get people to understand what I refer to as the authentic continuum.... But I think we need to get them to continue to go back and visit the "why" because we need them to very proactively plan for the strategy they are going to use and the authenticities so that it matches the "why of the evaluation. We want to get a shift from an evaluation that is simply for purposes of sorting to evaluation that gives students meaningful information so that they can improve. And again we want to emphasize that all our evaluations need to have self-assessments as the ultimate kind of goal. Then I want to work with them in terms of actually acquiring these exemplars that I'm talking about. (GI2, 9-94)

In an interview later in the year, he explained that he wanted the teachers to set their own goals:

Eventually, I want a staff that is designing and doing their own professional development. I should facilitate that, but this shouldn't be something that you do to them every once in awhile. If it's really working, they're doing their own self-directed, ongoing professional development and we think a teachers' professional portfolio will support that process all the time. (GI3, 4-95)

Dr. Getz explained that the portfolio will contain each teacher's own work plan which establishes his or her personal professional goals, the steps for reaching those goals, and eventually a self reflection of whether those goals have been attained, as well as a
list of professional development activities. Other items that the
teachers will need to include in their portfolios are:

They have to cite two to five examples of student exemplary
work and they can include it and they have to explain why
it is exemplary. They will use their portfolios for two things:
(1) self assessment, a vehicle for reflection, and (2) for
sharing. They will use their portfolio to create a showcase
portfolio that at the end of the year will be shared through a
poster session with everybody else on the staff. (GI3, 4-95)

When asked in June, 1995 how they felt about the portfolios, the
senior team teachers indicated that they had not heard about
them yet, but that they would probably find out about them at
the 1995 fall staff meeting.

Discussion

Was the staff development consistent with what the
literature suggests? There were a few mistakes made along the
way, but most of the plans, as far as they go, have followed the
suggestions of the experts. The administrators have been, as
Fullan (1991) suggests, the ones setting the tone for change (p.
191) and the initiators of new programs as well as the source of
support and advocacy (p. 54).

Dillion-Peterson (1990) suggests involving teachers in all
phases and respecting the wants and needs of individual staff
members (p. 224). The teachers were not required to attend or
take the workshop and/or assessment course. Initially, the
administrators made the mistake of planning the course without
involving the teachers, resulting in few teachers signing up because many had previous plans, and so, the course had to be postponed. But then, the teachers were brought into the process and asked what would be a convenient time for them. The course was offered free to the teachers using fee waivers, it was given at a convenient time and place (the school and right after the school day), and the teachers received college credit.

The teachers' wants were not taken into consideration as the teachers had neither input into the content of the assessment course nor into suggestions of other ways they might learn about authentic assessment. This is an example of the administration's failure to incorporate what is known about adult learning into the planning of staff development. Wood and Thompson (1993) claim that "adults want to be the origin of their own learning and will resist learning situations that they believe are an attack on their competence" (p. 55). Since the administrators' and the teachers' expectations of the course were different, and the course was based upon what the administrator's wanted the teachers to learn and the teachers felt they already knew that, resistance on the part of many of the teachers may have prevented them from hearing and learning the other information that was presented.

Following the teachers' formal introduction to the ideas of authentic assessment, the move has been to allow the staff development to become self-directed on the part of the individual teachers. This is consistent with what Lieberman (1995) and
Darling-Hammond and McLaughlin (1995) promote, that is, allowing the teachers to learn much the same way they are expecting the students to learn rather than something being mandated top down, and is consistent with a constructivist view of learning. Another suggestion of the experts is to provide sufficient time for reflection of the practices and time to plan. The central office administrators planned inservices following the 1993-1994 and the 1994-1995 school years as times for teachers to reflect, share, and plan. The senior team teachers welcomed this time and requested more.

Rewarding the implementation of practice is important to the success and continuation of change. The administrators have done nothing in the way of remuneration to the teachers beyond the paid inservices and pay increases based on so many college credits and masters degrees, but have been generous with praise and acknowledgments. They have encouraged the teachers to take risks, to visit other schools to talk about their program, and to attend conferences. The teachers' work has been rewarded by what their students have accomplished. The superintendent and the director of curriculum did not attend the first exhibition. But during the second year, they both attended the second exhibition, and the superintendent served as one of the judges. They recognized the teachers' hard work, and acknowledged the success of the Senior Seminar course and the second exhibition with judicious praise.
The teachers in the senior team beamed as they shared the words of the superintendent and the director of curriculum with me following the exhibition, "Dr. Getz told me personally to convey to the rest of the group that he thought exhibition night was wonderful, very well done." And then to these teachers who had perceived that they were always being compared to Sherwood, the ultimate praise was, "As a matter of fact when we went over there (to the superintendent's office), he was pleased with the way we were doing them (the projects), too, because he specifically said that it wasn't driven by technical writing like the other ones are, like at Sherwood, but that this was a free choice of the students that they could explore anything they want to do, and there is a writing component, but it is not based strictly on writing." "He thought that was a strength." (TI3, 5-95) They also mentioned that Dr. Jones had given them a critique, not of criticisms, but of things to think about.

Other suggestions for administration from staff development literature are: "rethinking schedules, staffing patterns, and grouping arrangements to create blocks of time for teachers to work and learn together" (Darling-Hammond and McLaughlin, 1995, p. 601) and grouping teachers into teams. This is precisely what the high school principal has done. He grouped the teachers into teams and arranged the schedules so that teachers have individual planning time, team planning time, and next year they will also have departmental planning time. The principal seems to
have stayed in tune with the teachers and where they are and what they are dealing with, while trying to push them ahead too. Wood and Thompson (1993) suggest that the principal is a key to school improvement practices:

Successful improvement programs usually find principals working collaboratively with their faculties to establish goals, working with teachers to plan staff development, participating in learning with teachers, providing follow-up assistance, and serving as an advocate for teachers who are implementing more effective professional practices in their classrooms. (p. 54)

Fullan (with Stiegelbauer, 1991) lists two things that principals have done in effecting change:

They showed an active interest by spending time talking with teachers, planning, helping teachers get together, and being knowledgeable about what was happening. And they all figured out ways of reducing the amount of time spent on routine administrative matters; they made sure that change had an equal priority. (p. 168)

Mr. Kent is knowledgeable of the skills required of his role as principal and change agent, and actively practices them. He talks regularly with the teachers, often visiting in the lunchroom and other informal gatherings. His participation in the assessment course, as he said, was to serve as a role model and "to promote a sense of collaboration among the staff, among the members of the class, and have a forum for sharing among the levels." He encourages and praises the teachers for taking risks, often playing
the role of mediator and cheerleader. A good example of this was observed as he described the way the team had handled a problem that surfaced concerning grading:

But the way they solved the problem is what impressed me. They got together and said, "We've got a problem here," and there was no fighting, there was just professional discussion and there was, "Oh let's involve the kids," and they called in the kids and said, "Oops, we screwed up," what a human face on a school. "We made an error but we are going to make it right." That is a very teachable moment for the kids, to have the four very important adults in their life make a mistake and then admit it, but my feeling was that these people and it was a major philosophical rift that could have torn a team apart that wasn't as high functioning as this group. (TI3, 5-95)

One other element mentioned as important in the success of any reform is the acknowledgment and willingness to work with the school culture (Deal and Peterson, 1990). This may have been overlooked by the administration which originally came into the school with a preconceived agenda of what should transpire in the next few years, and may have contributed to the resentment and resistance on the part of the teachers as they felt that the administration did not know or appreciate their talents, their knowledge of current educational issues, and their capabilities. However, over time, an appreciation of the teachers and their knowledge and dedication was apparent. Dr. Getz's respect of the teachers was expressed as he spoke of how teachers were making important adjustments in the curriculum based on their observations and conversations with one another: 173
So we were just going around sharing how we plan, well one of them said I know that Mr. Stem has asked and feels that his kids don't do well enough at oral presentations so all year now at the eighth grade we have been trying to incorporate more oral presentations. So we're trying to create the sharing climate and the talking stuff where you allow it to happen you can't make it happen. And you don't design it in, you encourage it you do things that will help that.... So we are honoring how teachers are organized and how they plan and facilitating the objective alignment process.... (GI3, 4-95)

Communication is another element crucial to the success of any reform. "To the extent that the information flow is accurate, the problems of implementation get identified. This means that each individual's personal perceptions and concerns - the core of change - get aired" (Fullan with Stiegelbauer, 1991, p. 199). Fullan also claims that, "Educational change depends on what teachers do and think - it's as simple and complex as that" (p. 117). Fullans' two points reflect much of what occurred at New Albany during the introduction of authentic assessment (and many other changes) by district administrators. Initially, the decisions and plans were set into motion by central office with little input from the teachers, but over time, the communication between teachers and administrators has been evolving and there seems to be more interactions between the teachers and the administrators from central office.

Communication needs to go beyond teachers and administrators, however. Stiggins' district plan for assessment
literacy, mentioned earlier, also includes a communication component. He suggests, "planning and conducting a community relations campaign in support of high quality assessment" (Stiggins, 1995, p. 245). Was this done at New Albany? The district developed a *K-12 Program of Studies* (1995) for distribution to "describe where we currently are in our journey" toward "a world class school system designed to prepare students for success in a global economy" (p. 2). However, the booklet did not elaborate on assessment beyond listing a variety of techniques that would be used to assess students at the different levels, and there was no mention of authentic assessment. In the booklet's description of the Senior Seminar course, the formal preparation program for the Ohio Twelfth grade proficiency test was mentioned.

The teachers in Senior Seminar were concerned with communication and held a parents' night to explain the course, including assessment, but were careful not to mention anything that might result in debate. They also prepared a pamphlet that contained their philosophy which was given to parents and students.

So communication to the public, in terms of any campaign in support of high quality assessment or authentic assessment, has been delicately handled. The controversial nature of assessment and testing, the public's concern with their district's standardized
test scores, and some of the problems from places like Littleton, Colorado (Davis and Felknor, 1994; Shepard and Bliem, 1995) have many shying away from public discussions of this issue.

Another part of Stiggins' plan has to do with "setting policies that demand and support high quality assessment" (p. 245). Many of the policies have been addressed in New Albany's strategic plan. However, the district does not have any hiring policies regarding assessment literacy. Dr. Getz explained to me that the district has no written policy, but during the hiring process, the administrators look closely at what courses prospective teachers have taken and what they know about assessment. If a teacher has "no clue," in terms of assessment, he or she probably would not be hired. Once hired, however, new teachers become involved in a mentoring program to acquaint them with what New Albany is doing, including assessment. Unfortunately, Dr. Getz reports that the mentoring program is not working in the way it was planned. By setting specific policies regarding assessment in terms of hiring, New Albany could send a strong message to parents, students, and community members, as well as colleges and prospective teachers as to the importance of assessment and its role in education.
QUESTION 2: What are the attitudes of teachers toward assessment and what were their practices regarding assessment before, during, and after completion of an assessment course and/or other introduction to authentic assessment?

For many years in schools, the role of assessor has been delegated to psychometricians and separated from the role of teaching. Today, however, many believe that instruction and assessment go hand in hand, and that teachers must assume the primary role of assessor. Teachers have always been what Stiggins (1994) calls "the hidden assessors" (p. 33), those doing the day to day testing and assessing of students, but have received little if any assessment training. They are usually unfamiliar with the technical aspects of assessment such as reliability and validity. Of the four teachers in the senior team, one remembers learning about testing in his method's courses, another had some sort of testing course in his master's program, and the other two had statistics courses, but not a testing course per se.

Resnick and Resnick (1992) identify three main classes of educational testing; each class serves a different purpose and a different audience:

(1) Public Accountability and Program Evaluation
In this category, the results of the tests are used primarily by people removed from the day to day process of schooling and are
intended to provide information on the overall effectiveness of the programs rather than detailed information on students.

(2) Selection and Certification of Students
These tests determine selection of students into institutions or programs, and they should predict the individual's future performance in them. Examples are the SAT and the ACT.

(3) Instructional Management and Monitoring
This class is more pertinent to what goes on in the classroom; diagnosing strengths and weaknesses in student performance and understanding and monitoring student progress towards specific objectives. According to Resnick and Resnick (1992), "for instructional management tests to work effectively, they must be tied directly to the curriculum in use, providing information about exactly what aspects of a syllabus individual students know and do not know and which problem-solving and reasoning processes they can or cannot manage" (p. 51).

Teachers' Views of Assessment

Prior to the assessment course, the teachers in the senior team were asked, What do you see as the purpose of assessment? They described their purposes in diagnostic and monitoring terms. Ms. Dearheart answered in an interview in September, 1993:

I think teaching is a constant assessment, and I have to take a look at every question I ask in class and what kind of
response did I get, not only from the student who answered out loud, but from the other students in the class.... The role of assessment is huge. You can't teach without assessment, they go hand in hand. If I just stood in front of the room and lectured and paid no attention to how the students were processing the information that I was wanting them to learn, there would be no reason to teach. You have to recognize the individual. (DI1, 9-93)

Researcher: "So you're diagnosing all the time?"

Ms. Little answered:

Absolutely, good word. And that's part of what I showed you the other day, that's why I'm writing those things down especially for that other class of 25 because I can't keep all of that in my head, and so as I go on through this, I'm finding out the ones who aren't doing their homework and if there is any correlation and are their test scores going to be lower if they haven't done it or are they one of those who it isn't necessary to do their homework, I've had one or two of those. So, yes, it's constant, a constant process. (DI1, 9-93)

Researcher: "In the past, what kind of tests have you done, has it primarily been multiple-choice, word problems?"

Ms. Dearheart responded:

I have never, how shall I say, I very rarely have given multiple-choice, true-false, fill-in the blank, almost, almost never. On rare occasions I will do something like that. Most of my tests are what the AP [Advanced Placement] terms as free response. I give them a problem and you do the problem and show me how you did it ... you tell me how this problem was done so that I can see your thought process, and that's why I'm so particular about, especially about the beginning of the year telling the kids, I need to see your work.... understanding the process is different than making an operational error, very different. (DI1,9-93)
Mr. Stern described the purpose of assessment:

To measure growth, to try to determine what, I hate to say the word learn, ah, to get a picture of the student in other words, at a particular time and place. I don't think it's necessarily, you know everyone says learn, what does learn mean? Learn means different things to different people. I know what the traditional definition of learn means, sort of a paper and pencil test and measurement by standardized tests and such. There's a lot of ways to learn. I like to judge the students in as many ways as I can. I don't like just one picture of them. For instance, in my class, testing is not that important. I don't necessarily think paper and pencil tests, as a matter of fact, I don't even like them. I'd say that 20-30, maybe 40% of the grade is all other measures. It's easy. That's why people use it. It's a simple way to, it's an easier way to do things because it's pat. You make the test up, you give them the test, you grade the test and say, "Here you are." I think that's only one way to measure things. I like projects. I like simulations. I like reports. I like homework. I grade that as much as I do tests. So I give them as much point value for that so a student who doesn't necessarily do well on testing, and I don't believe that testing is very meaningful in terms of longevity, I think it's mostly short term memory. I think it's just like everybody does, you study when you need to study, you do what you need to do and then you're finished, through, and you promptly forget everything. You remember some of the things, but unless it is really meaningful to you - I interviewed my seniors last year in my government section and none of them, not one of them, said, "One of the best things this year was testing." (SI1, 1-94)

Ms. Little explained in an interview that the purpose of assessment for her was:

a better evaluation of student progress and development: personal and intellectual, in many different facets really. I have used over the years probably mostly paper and pencil
since that is how I was taught and I guess we are most comfortable with that. But I don't know how long ago, probably with learning styles other than visual, I am a very visual person, so that really started it, so I don't know how long ago, but I have continued to try and develop other ways and to tap into strengths that the students have other than just paper and pencil. I don't give major tests anymore. I give lots of quizzes to insure that students read. Otherwise I feel they probably wouldn't. So they have quizzes and there is always an essay on there. There is multiple-choice, usually, and then there is always an essay. We do journal writing. We do a fair amount of that. I give projects. The students don't typically like that because some people end up doing the work and everybody gets the same grade.... life is not a matching or it is not multiple-choice, it is not a test like that. It's more, "How can they apply it and still demonstrate that they understand it and they have integrated it. And that's a hard thing to do. It is much easier to grade a true-false or a multiple-choice or those kinds of assessments. (LI1, 1-94)

In Mr. Bradford's interview before the course started, he explained the purpose of assessment "to evaluate the progress of the student through the curriculum." Methods he used were one-on-one evaluations of progress, tests, quizzes, oral presentations, and video production. "Basically I try to do just about everything that I've heard of one time or another. I find some to be more effective than others in certain circumstances." (BI1, 1-94)

All four of the teachers perceived the purpose of assessment in terms of determining where the students were and as a means of providing necessary and useful feedback to the students. It was not viewed as a sorting or selection tool or in a punitive way. The methods used to diagnose and monitor their students were
numerous and varied. They used a combination of objective measures as well as alternative assessments, but they were all interested in learning new ways of assessing the students and expanding their repertoire of tools.

As part of the assessment course assignment, the senior team created a philosophy statement that captured their thoughts and practices which they refined and modified to share with students, parents, and others (see Appendix C).

What about the teachers' assessment practices in the Senior Seminar course? What did they do early in the first year? Did they change assessment strategies and practices, during and after the course? Did the teachers incorporate what was taught in the assessment course into the Senior Seminar assessments?

When the Senior Seminar course began in 1993, the team teachers were working with an ungraded course, and therefore, the students received their grades for work done in the course through other courses. There were no scoring rubrics or written criteria for the students to follow. At this point, as the teachers have said, they were learning, too. They were familiar with authentic assessments, for instance, Mr. Stern had attended the assessment workshop that summer and was using several authentic activities in his classes. Ms. Dearheart was incorporating portfolios into her Algebra class. But there were so many other concerns with just getting the Senior Seminar off the ground the first year and there were no guidelines, so the teachers had not
yet crystallized the assessments that would be used in the course. The teachers knew they wanted to provide meaningful activities for the students. They would allow the students to choose the topic of their project as well as the format by which they would present that project, and there would be an open exhibition at the end of the year for the top projects. The teachers also knew that they wanted the students to create a portfolio at the end of the year that would contain items that showcased what the students had accomplished, something which could be used for future employment or college entrance.

The assessment course for the teachers coincided with the period that the students were working on their projects. The timing was perfect because it allowed the teachers the opportunity to think about the assessments for the projects and Senior Seminar in a supportive environment. As part of the assessment course assignments, the team created their philosophy and assessments for the projects and presentations. As they noted in an interview:

It [the course] probably didn't help us in terms of giving us a lot more information, although we did know some of those things and some of it we didn't. I think what it did was jump start us. It gave us a reason to sit down and start planning. (TII, 5-94)

They also developed the criteria and scoring rubrics for the projects and presentations during the assessment course. Following the course, the professor had offered the teachers, and
they accepted, the option of taking an independent study with him in which they would implement the assessment plan they had created, and then critique it: what worked, what didn't, what would they do differently. The teachers presented the scoring rubrics to the students a couple of weeks before the projects were due, and then used these to grade the projects and presentations. Ms. Little reported that when the rubrics were presented, one student had commented, "It's a little late!" She explained to the students that the rubric would be useful for the things they still had to do, like the presentations. She confessed that the teachers were also learning, along with the students. The scoring rubrics were also given to the judges for the exhibition, but with no instruction or guidance on how to use them. The self-assessment component of the project consisted of having the students write an evaluation of their effort on both the process of creating the project and the product itself. As part of the independent study, the teachers also worked on the indicators for the portfolio which had to include at least six items from a checklist.

The implementation component of the independent study assignment occurred during the last two months of school, but the teachers had not had a chance to sit down and critique it yet. They were able to do that during the district inservice following the 1993-1994 school year and several days beyond on their own time. The team teachers reflected on what had worked in Senior Seminar and what had not, they then modified and revised the
schedule, the activities, and the assessments for the following year. They also worked on the objectives to be included in the graded course of study.

The things done the first year would probably not qualify as fair assessments. Since the teachers were not clear about what they wanted the students to do, the students did not know what they were supposed to do. But the teachers used what they learned from the first year and what they saw to be strengths and weaknesses in what students knew and could do. They saw that the students needed to develop better speaking and presentational skills, and needed further work in organizing and management. They built these into the curriculum and the assessments not only for their course, but shared them with teachers in the earlier grades, and these became part of the 9th grade curriculum.

**Changes for Second Year**

In the second year of Senior Seminar, the course became more structured with specific plans, clear expectations, criteria, and rubrics:

Mr. Stern: even a couple of the things we really planned well on last year that were good we still needed a lot more impact on the students and a, the skills were not being done anywhere yet, they are still not being done in the lower levels although some people are trying; we have to do that, there just isn’t anyway around it, we have to help them to learn these things so they can do a better job of the things we want them to be able to do.
Ms. Little: We determined that a year long project was too much, that the kids cannot focus that long on one project, so that was kind of like probably the driving thing, to pare that down to nine weeks and then to organize the rest of the year into quarters, and have something for each quarter which we think has given a lot more structure and direction for the kids. Although, ironically, I have had some people say to me, I wish these projects were year long. (Laughter)

Ms. Dearheart: One of the biggest factors was that the course became a credit and a requirement for graduation and so we had to assign grades and come up with what we thought was fair and equitable. ... Consequently the rubrics I showed you and being upfront with the kids and saying, "This is what we are going to do and this is how you are going to be graded," and I think they have a more clear understanding exactly what it is we expect of them because they know upfront what it is we are going to grade them on.

Researcher: Did the rubrics become more specific?

Mr. Stern: We have one for everything; in other words, just about everything we do, we have a rubric for it; they haven't changed that much.

Ms. Little: No; maybe perfected a little bit, but the essence—very, very minimal changes on the rubrics.

Ms. Dearheart: Like I indicated before, what we are trying to do is to make them all uniform across the four nine weeks, so that the kids are very used to them by the time they get to the project, they know exactly what it is they will be graded on when they stand up and give a presentation. And this is our speech rubric. (T12, 12-94)

Student expectations were very clear. The teachers had talked to the juniors, before the end of the 1993-1994 school year, about Senior Seminar and had shown video clips of the
projects and presentations that had just been presented. So these
students already had an idea of what would be expected of them
when school started in the fall. At the beginning of the 1994-
1995 school year, the teachers explained to the students in detail
what they would be expected to do this year and showed them
exemplars of the journal entries and the portfolios from the
previous year. The teachers provided course materials which
outlined the course philosophy, goals, the four strands, ground
rules, and an explanation of the assessments that would be used.
The sheet on assessment read:

Assessment is intended to provide useful feedback to
teachers, students, parents and others interested in real
student achievement. Students will be given opportunities
to complete a variety of performance tasks demonstrating
their understanding and application of knowledge, skills,
creative thinking, decision making, problem solving, self
management and collaboration. Assessment will often occur
over an extended time, result in tangible product or
observable performance, depend upon individual and group
work and result in demonstration to an expanded audience.
Using established criteria and standards which reveal
degrees of proficiency, assessment in Senior Seminar will be
made through the following tasks: journal entries, portfolio
development, public speaking, mentoring/shadowing
experience, mock interview, short term experiences, long
term projects, community service experiences, self
assessment, twelfth grade proficiency test.
Senior Seminar is a one credit requirement for graduation
from New Albany High School. All work assigned will be
evaluated as 100%, 90% or 'Work in Progress'. All work
evaluated as 'Work in Progress' must be satisfactorily
completed two weeks after the end of the grading period the
work was assigned or that work will earn a grade of zero
(0%).
The teachers also included what assessments would be used for each quarter and the percentages that would constitute their semester exam grades. Their written explanations were verbally reinforced in the first few days of the class and reinforced throughout the year. They tried to make it very clear to the students what they were expected to do.

A meeting was held with the parents the second week of school to explain the purpose of the course and what would be expected of the students. According to research (Shepard and Bliem, 1995), educating parents about the new forms of assessment and bringing them "on board" has been an important issue in the success or failure of programs.

According to Jones and others (1994), when using authentic tasks to promote engaged learning, it is helpful if students are "involved in the generation and evaluation of the criteria used" (p. 12). Dr. Lucas emphasized this idea in the assessment course and demonstrated how this could be done with students by having the teachers generate the criteria to be used to evaluate their own assignments. In the first year, the teachers created the criteria and rubric for the projects and presentations a week before they were due, and the students complained that it was given to them so late. Upon reflection at the end of the year in a team interview (TI1, 5-94), the teachers discussed the criteria and rubrics, and Ms. Dearheart suggested that they involve the students next year, "I think we need to sit down with the kids and have them help us
develop the rubrics for next year. Most of the rubrics, we kind of bounced things off them [the students], but I think we need to get a little more specific than that and ask the kids to develop what it is they want us to look at." Ms. Little offered, "They don't know though, if you just say, 'what would you like us to look at'." Ms. Dearheart: "No, no, we need to lead them. But I think that they need more involvement." But this did not happen. Instead, the teachers refined the criteria and rubrics without input from the students.

After the second year, when I asked the teachers if they planned to involve the students for the third year, they expressed concern about the time that would take. Ms. Dearheart responded, "I think that it is a good idea, but trying to manage 60 kids and get them all thinking about the same kinds of things, thinking along the same line at one time is a tough job." Ms. Little continued:

And I don't think kids are very good at that at all, maybe it's because they haven't been given the opportunity or they haven't even started to think about those things, but that would be a real leading experience. I think we would have to feed them, not to give us the answers we want, but they haven't even thought about stuff like that let alone vocalize it or verbalize it. (TI3, 5-95)

Although the teachers did not involve the students in creating the criteria and rubrics, they did develop more rubrics, and revised and modified the old ones for the 1995 projects and presentations, making them clearer and more well defined than
the previous year's. This is consistent with what the literature suggests. Wiggins (1993a), Arter (1993), and Herman, Aschbacher, and Winters (1992) emphasize the importance of making criteria crystal clear for students. Arter (1993) claims that, "Whenever we make a judgment about anything we use criteria whether we can articulate them or not" (p. 1). It is only fair, then, that we clarify the criteria up front rather than make students guess what it is we, the assessors, expect. Wiggins (1993a), however, notes that "not only is diversity of response to be expected; the correct answer may not even be specifiable" (p. 10), so it is the responsibility of the assessor to "develop apt scoring criteria and standards that reward diverse excellence" (p. 11).

The teachers' growth in assessment savvy and experience in grading alternative assessments with several scorers is illustrated in a problem they encountered and how they dealt with it which they described in an interview in the spring of 1995:

Ms. Little: There was a real problem with one of the rubrics this year which Mr. Stern and I were misinterpreting.

Ms. Dearheart: No, We just miscommunicated.

Ms. Little: So we said, "Oh wow, we need to do something about this." But we need to do something permanently about it for next year. It had to do with

Mr. Bradford: the super team presentations.

Ms. Little: Somehow converting some scale to some scale.
Ms. Dearheart: What was happening was Ms. Little and Mr. Stern were thinking -we had 0,1,2,3,4 scale- and they thought 4 was an A and 3 was a B, but Mr. Bradford and I were thinking in percentages because if you gave them a 3 that was 3/4's which is 75% which is a C so we had to

Mr. Stern: When we looked at the projects we had a different-

Ms. Little: We thought we were giving them Cs and we were giving them Fs.

Ms. Dearheart: So consequently our scores were a lot lower than we thought they were going to be.

Mr. Stern: So we had to call them (the students) all in and kind of apologize to them and work it out somehow.

Ms. Dearheart: so we decided on ours [project rubric] we would go strictly on percentages, ... if we gave them 10 out of 20 points, then that was a 50%, so that we were all thinking on the same wave length.

Ms. Little: Ms. Dearheart did a whole scale.

Mr. Bradford: We didn't grade them in common. We graded them, compiled individual numbers. (TI3, 5-95)

In the first year, each of the teachers only watched half of the presentations, but in the second year, all four of them watched all of the presentations. The change was made for two reasons: first, the teachers wanted to see all of the presentations and second, they felt it wasn't fair:

Ms. Little: I was sorry I didn't get to see some of them, to the fact that maybe it's not real fair unless all of us watch all of them. (TI3, 5-95)
During the first year, the teachers discovered that each one graded a little differently, and even though they had felt that having a "harder" and an "easier" grader on each team averaged out the scores, by the second year, they realized they needed more consistency.

The scoring rubric created for the judges the second year was also quite different from the previous year's, and enabled the judges to be more consistent and much more efficient:

Ms. Dearheart: One judge [who had also been a judge the first year] said to me he appreciated the numerical system and that we tallied the scores and that they didn't have any part in the deliberation. (TI3, 5-95)

During the first exhibition, the judges were not given any instructions on how to execute the scoring rubric, and the deliberation process took forty-five minutes. Afterwards, the teachers had no idea how the judging had been conducted. So, for the second exhibition, the teachers refined the rubric, sent it to the judges ahead of time with instructions that they were to mark the scores as the students presented, and then the scores would be collected at the conclusion of the exhibition, tallied on the computer, and the winners selected on the basis of the total scores. Although, the teachers have worked to make the scoring rubrics more reliable by defining them more clearly, and by talking and discussing what they mean by the various scores, this is still an area that the teachers should continue to revisit.
In addition to more structure and making the expectations very clear to the students, the senior team instituted the concept of "redos." Upon reflection on the projects produced by the students in the first year, the teachers felt that, by refusing to give anything below an A or B, and requiring the students to redo unacceptable or substandard work, the students would try harder and the overall quality of the projects would be improved. This is consistent with the idea that all students can learn and that the purpose of assessment is to provide feedback and scaffolding. According to McDonald (1993) it is a matter of equity in assessment; "how to get everybody to a place where some never imagined that they could get" (p. 485). The idea of redos is also consistent with Wiggin's contention that "performance is only effective and meaningful when work is reworked" (Wiggins, 1995a, p. 105).

During the second year, the teachers felt that the students' products were of better quality than the first year. Did the redos contribute to the better quality work? When the year began, the students had received a copy of the prize winning resume, had been shown the prize winning process journal, and had seen the videotape of the finalist's projects as exemplars. The teachers explained:

Ms. Little: I don't know if that [the exemplars] had anything to do with the quality. It's simply that we wouldn't accept it until it was done right. It kept going back and going back and going back and they never got any credit until it was done picture perfect and we deemed it was almost perfect.
I guess perfect, not almost, it had to be right. It is either right or we are not taking it.

Ms. Dearheart: And there are some things throughout the year we have done that way. The project update-either this was done correctly and we knew exactly what their project was about or I gave it back to them and said "work on this some more." (TI2, 12-94)

But the teachers became overwhelmed with the amount of grading for the journals and the other activities. By November, their ambitious goal of having the students receive nothing below a B by having them redo their work had dwindled to only the major assignments.

Grading changed dramatically for the second year, primarily because Senior Seminar became a graded course of study. Two components went into the students' Senior Seminar first semester exam grades for the 1994-1995 school year: 50% was based on project development and 50% on community service. The second semester exam grades were derived from their scores on the state's senior proficiency test. According to Mr. Kent, the 1993-1994 senior class didn't take the test seriously, and only 48% of the students passed all four parts:

I'm trying to remember; very, very few passed any sections at all with honors, they basically played connect the dots; .... This year the percentage went to 73% passed all four .... I had one student pass all four with honors, ... now the difference is, last year's class was an excellent group of kids, this year's class is a great group of kids. A couple of things we did: we tied senior proficiency test scores to Senior Seminar; their senior proficiency results will become their
second semester exam grade in Senior Seminar. I personally met with the seniors several times and we're in the process right now of negotiating their reward for doing so well. We established a target of 60% of the class passing all four and they exceeded that by 13 points; so you know they just worked hard at it; they worked harder than the kids did last year. Also the senior teachers in Senior Seminar did a two week preparation in the Senior Seminar period ... in content stuff, test taking, the kids that we knew, for example. that hadn't had a math class for two years did some review sheets and went through some stuff. It wasn't that intensive, but that is one of the things we are going to work on for next year is sort of do kind of like an SAT prep class. (KI2, 4-95)

The teachers felt that this strategy of using the student's grade on the proficiency test as part of the Senior Seminar course might provide more motivation for them to take the test seriously. They also spent a couple of weeks getting students ready for these standardized tests. The teachers had to deal with the district's push for more authentic assessment and the state's push for minimum competency proficiency tests and the Senior Proficiency test, as well as many other objective tests that the students take such as SATs and ACTs.

**Administrators’ reactions to assessment changes**

The adjustments and changes that the teachers made in the Senior Seminar assessments for the second year were consistent with the expectations of the administration. Dr. Getz shared his thoughts about standards in an interview:
They had the exemplars from the first year to share with the kids, .... One of the ways we think you can address improving standards is by collecting and sharing exemplars. Second way is having your rubrics, having clear criteria that are communicated and not secret. Third way is have precise language when you're generating your instructional objectives and goals, language that is very explicit in terms of the kinds of performances you will expect. (GI3, 4-95)

Prior to the second exhibition, Dr. Jones voiced his approval of the team's assessments:

I just saw a rubric. I'm one of the judges at the end of this month. I'm pleased, that's a work in progress and it needs to be retended every year and there will be new dimensions added to it.... We need to get community support, those almost become my personal rubric for the program at large. I trust entirely, Mr. Stern and the professionals in the class to set up the rubric for the kids within certain parameters. I just want to see a lot of writing skills and communication skills in evidence, the standards need to be rigorous. How that is mapped out within that umbrella is their job and I'm very proud of them. (JI1, 4-95)

**Discussion**

The team's whole approach to Senior Seminar follows Gardner's (1992) "assessment view" in contrast to the "formal standard testing view:"

The "assessment view" values the development of productive and reflective skills, cultivated in long-term projects. The animating impulse seeks to bridge the gap between school activities and activities after school, with the thought that the same habits of mind and discipline can be useful in both kinds of undertakings. Especial attention is paid to individual strengths. (p. 115)
The assessments created for the course during the second year and the changes made were consistent with what the literature suggests for performance and authentic assessments: making the activities meaningful for the students and their future; providing the students with exemplars of good and bad work; providing criteria and rubrics to the students in advance; increasing interrater reliability with scoring rubrics; informing and educating the parents; providing motivation to the students by allowing for practice and feedback and letting them know that by working hard and, redoing when necessary, they can attain a good grade; expanding the audience beyond teachers and peers; and a self-assessment component. The teachers use a variety of assessment tools to gather information on each student's process and product: portfolios, journals, interviews, project, presentation, and self-assessment. Although the things that have been done are positive and consistent with authentic assessments, many are initial attempts and need to be taken to the next level.

One component which the teachers had not yet incorporated into the authentic assessments was that they have yet to involve the students in the process of creating the criteria and rubrics. Allowing students to participate in this process, gives students ownership of their work and helps define clearly for them what they are expected to do. Perhaps as the teachers get the course more finely tuned and deal with some of the other issues, they may have more time and the confidence to allow the students to
participate in this process. However, there is also the possibility that as the teachers get the course and assessments more refined and become satisfied with how it runs, they may not want to take the risk of involving the students and have the assessments change dramatically. Having come to know these four, knowledgeable, caring, professional teachers, I would expect the former.

The scoring rubrics need to become more defined and explicit. This is especially important for parents and the exhibition judges. The teachers and students may have come to agreement on what constitutes a score of "one" or a score of "ten" through exemplars, oral discussions, and daily interactions, but parents and the judges do not have these same advantages and opportunities, and thus need better guidance and clarity regarding the scoring rubrics. The community at large probably exhibits a wide array of perceptions on what constitutes a "one." More clearly defined indicators would provide better consistency in grading and would be fairer to the students.

Within the two years of this study, the teachers did not use self-assessment and peer-assessment to their full advantage. Although students were expected to self-assess their work when it was completed, they had not been adequately prepared to do so, and therefore, were not using the criteria and scoring rubrics to assess, for themselves, how they were doing during the process.
and then making revisions accordingly. One of the second year students explained how he used the criteria and scoring rubrics:

They gave it to us, it was in our notebooks, and it was in there, and I looked at it once in awhile. And I was like "Oh well, I'll just kind of wing it, I'm better at living than at preparing things like that," And so I just looked at it "Oh OK," and I kind of understood, but I didn't want to grade myself or anything like that, that's just not the way I do things. (SI, 5-95)

Perhaps, if the teachers work with the students to help them learn how to self-assess, and provide opportunities for the students to practice self-assessment and peer-assessment, not only will this benefit the students' future learning endeavors, but it might also help with the teachers' time problem. If students assess their own work as they go along and before turning it in, there should be fewer mistakes, and consequently, fewer redos.

Another area of discussion in relation to alternative assessments has to do with the difference between standards and standardization. Traditional assessments require standardization to assure that everyone is treated the same way, has the same questions, are given the same amount of time, and everything is kept secretive so that no one has an advantage over anyone else. This served a purpose given the behaviorist model of learning which assumed that the teacher had all the knowledge and could pass it on to the students and they would all learn the same thing. However, given a constructivist model, standardization is not
appropriate. Standards are not the same as standardization, but play an important role in authentic assessment and in educational reform. The team teachers, in some cases, seemed to have confused the two concepts.

Their purchase of the journals so that everyone could write the same amount for a page is an example of standardization. Their insistence that everyone write two journal entries every week is again standardization. Commenting on the weekly journal entry about her project, one student complained:

'It got kind of frustrating on what to write, because some weeks I would do a lot of things, and on my own I would document it in just notes, as far as what I had already done and what I needed to do. But some weeks, you would just have to sit there and just kind of come up with stuff to write, like little things. So I don't really agree with it having to be on a weekly basis. I think you should just be on your own time as part of your project to turn in just the documentation of what you do. If you do something, write it down for your own reference. That way you can keep track of your own project. I think that would be more beneficial. (SHI1, 5-95)

The standards, on the other hand, are still in the process of being identified and discussed by the teachers and administrators, and this process should be extended to the larger community.

Did the teachers incorporate authentic assessments into their other classes or did they use it only in the Senior Seminar course? I asked the teachers, "Have the assessment practices changed in your individual classrooms?"
Mr. Stern: In my government class, I use more rubrics now for grading everything, it's easier, I've found it to be much easier than what I did before and in AP I do rewrites on those. I really need to do more community kinds of things that would be a valuable thing for projects evaluations or something like that; that might be something I will try to do for next year.

Ms. Dearheart: I find myself looking for more open ended kinds of problems, so that the kids have more flexibility in the solutions of problems and that gives the entire class a broader picture of the topic.

Ms. Little: How could it not? (TI3, 5-95)

The whole idea of Senior Seminar and its assessments are consistent with the concept of curricular coherence. Wiggins (1995) views a culminating senior experience, such as the Senior Seminar, as the last of many twists to the school spiral which allows students to build and pull together their knowledge and understanding. He suggests that the assessments required in such a curriculum need to be more longitudinal:

We will want to chart novices' performance gains on a continuum that ends in expert performance.... Curriculums would then always be designed backwards from expert, fluent performance and genres that we want novices to master; influencing even the work of curriculum design in the earliest grades. We would then seek what yields the most likely progress on the continuum, designing in the earlier grades and courses "milepost" performance requirements that are simplified "scaffolded" versions of final performances - as opposed to overly simplified, decontextualized lessons. (Wiggins, 1995, p. 117-118)
The district seems to be working toward this coherent curriculum as teachers and administrators have looked at what the students have done in the Senior Seminar and discussed how this should effect what goes on in the earlier grades. They have introduced two major items into the ninth grade language arts program. Dr. Getz explained:

one being research related to information access and organization, and the second one is presentation skills.... not just oral presentations, but all the technologies that might go with that, whether it is video or computer programs or whatever. (GI2, 8-94)

McDonald (1993) discusses the concept of exhibitions and presents several issues which should be addressed. He explains that exhibitions should "have attachments" between students and teachers, but should be "balanced by cool judgments." Is this the case at New Albany? There is definitely a "warm attachment" which develops between the teachers and students as evidenced in the first year when the students signed a petition for the teachers to hand out the diplomas and the strong emotions displayed at graduation.

McDonald (1993) explains that "cool judgment" can be maintained by going back and looking at the videos of the presentations rather than totally relying on the judgments at the time of the presentation. The senior team teachers did their grading at the time of the presentations, but looked at the videos later to select good and bad exemplars for the next year's
students. McDonald also contends that the exhibition should effect
the curriculum. As just discussed, the results of the projects and
presentations at New Albany are having an impact on the
curriculum.

The first two groups of seniors to be required to take Senior
Seminar were perhaps at an unfair disadvantage. Since the
projects and presentations were to be done through an integration
of their work throughout their school years, it might have been
better to have started with a group of freshmen and explain to
them what they would be expected to do as seniors, as was done
at Central Park East Secondary School for their Senior Institute.
They added a grade each year as the first group of 7th graders
progressed through the system (Darling-Hammond, Ancess, and
Falk, 1995).

QUESTION 3: What problems, frustrations and concerns
have arisen with the introduction and implementation of
authentic assessment into the district and the classroom?

With changes, inevitable issues; concerns; and problems
arise. The teachers encountered similar problems and frustrations
to those that have been documented in the literature: lack of time
to plan, implement, and grade authentic assessments; problems
with reliability when using multiple graders; the lack of models
and guidelines; lack of equipment and extra people to help with
the paperwork and routine work; and problems with lack of knowledge about authentic assessments on the part of the students and parents. The issue which seems to be less well documented is the frustration and conflicts that inevitably arise when people in various roles who have different perceptions, expectations, and ideas about what and how things should be done try to work together. All of these are addressed below.

**Time**

What has been the biggest concern or problem encountered when trying to create and use alternative assessments? I asked. Without any hesitation, the teachers responded, "Time!" "Yes, keeping up."

Lack of time is acknowledged from all perspectives as a very real problem and a difficult one to arbitrate. Mr. Kent, the high school principal discussed the problem:

The time is the biggest issue and there are only two resources we can give them: time and money, and since we can't give them any more money because we are real structured on how we can do that, but the time is the one thing we can create for people and the common planning time is the single most important thing we did this year that was able to promote that for people in a situation where they can succeed. (KI1, 4-94)

Lack of time is an issue that has come up again and again in reference to alternative assessments. There never seems to be enough time for teachers to get everything done, but alternative
assessments require even more time to plan, implement, and grade than objective tests. In an interview with the senior team at the conclusion of the first year, the teachers described the time it takes to develop these authentic assessments as "phenomenal." The fact that they were working as a team also required more time:

"It's not like working by yourself, because if I'm working by myself I just cut the corners and I stick this in here and put that in there. Now it's negotiations. It's, we're bartering for time all the time, we're fighting for every minute." (TI1, 5-94)

Half way through the second year, they conveyed similar problems when discussing their institution of the requirement that students redo unacceptable assignments. Initially, everything was to be done this way; that, however, soon proved to be too time intensive and by mid year Mr. Stern concluded that:

"We believe that the most important things need to be done that way; we were getting swamped, we were getting bogged down in paper and we're not even sure whether that is even doable as we envisioned it. It may have to be modified to maybe just certain important things like the resume and project reports and things like; the things that absolutely have to be done like that but other things we prefer them to take zeros (laugh)" (TI2, 12-94)

Ms. Little: Because you can never get rid of them then; they are just with you and with you, you grade it a zillion times; we don't have time to do that anymore. It was good in theory, it wasn't so good in practice. (TI2, 12-94)
Mr. Bradford explained after the second year:

We found it difficult to have the students do the projects over if they were substandard. There just wasn't time, and we may have to come to some sort of compromise there in doing that. (TI3, 5-95)

By the end of the second year, their ideas about redos had expanded beyond their own course. These teachers thought that if redos could be initiated at the lower grades, eventually, by the time the students were seniors they would be accustomed to redos and the expectation of quality work, and consequently, the teacher workload of grading so many redos would hopefully be reduced:

Ms. Little: The redo issue is a powerful instrument...

Principal: How is it powerful, to you or the kids?

Ms. Little: To the kids.

Ms. Dearheart: By the time we get them [the students], there should be very few redos. They should know up front we want quality work the first time around.

Ms. Little: There was a point where we as teachers said if we got anything we were happier than if we got nothing.

Mr. Stern: Nothing is an alternative, too.

Ms. Little Yes, nothing is an alternative for them. So if we got something, something is better than nothing, but I think that if quality is the buzz word here, then we need to seriously, as an entire faculty, do that; redo, redo, redo.

Principal: You have the students redo it until it's right, acceptable, and I think that needs to become a philosophical
shift. How do you deal with the frustration of getting either nothing or substandard work again, substandard work? and how long do you let them redo it over and over, where is the bottom line? when do you say this is it?

Ms. Little: And, how do you score that? Plus, just keeping track of all that.

Principal: No English teacher should have more than 60 kids a day. There are reasons to have that. But if you are going to have those kinds of expectations, you better not have. We have to have more English teachers or whatever. That's a way to expend resources; you've got to have support on the board of education and I think we will. (TI3, 5-95)

From this discussion among the senior team teachers and the principal, some powerful ideas were emerging, problems were being considered, and possible solutions contemplated.

The projected New Albany school enrollment in the year 2004 is "between 2,000 and 2,400 students - an increase of 125 to 170 percent above the current enrollment" (Building a dream, 1995, p. 74). As the number of seniors increases, the problem of grading will grow proportionately. The senior team has already encountered this problem as their Senior Seminar class nearly doubled in the second year. The teachers discussed ideas the team had concerning how to deal with the larger number of students:

Mr. Stern: We've been talking about getting some help either with another faculty member or maybe somebody just to do the paper work because there's just a whole lot to do here: constant grades. (TI2, 12-94)
Ms. Little: Except that then we think about adding somebody to the team and saying, "it already takes the four of us X amount of time to do something and everytime you add somebody, it is going to geometrically grow and be more time and more time and more time." It takes a lot of time to plan. (TI2, 12-94)

Mr. Stern discussed the time issue when asked about the grading of the authentic assessments, like the process journals and the self-assessments:

We grade on our own time. That is something we can use help with or that we could have more time with. We'd probably suck it up with planning, so we'd end up with not as much time for grading in the first place anyway. Yes, we could always use more time. (TI2, 12-94)

The teachers also talked about the time necessary to grade all the senior projects and presentations following the second exhibition:

Ms. Little: It takes an inordinate amount of time to have all of us watch all of them, but we didn't like the fact last year that we only got to see half of them ... maybe it's not real fair unless all of us watch all of them, but oh my stars, the time involved. (TI3, 5-95)

The teachers discussed how they might handle the amount of time involved in grading all the presentations with the principal. Mr. Stern suggested restructuring the time by reallocating the fourth quarter's time away from community service which they decided could be "ongoing" throughout the year. That would free up more time for the projects and the
grading of them. Mr. Kent suggested stronger enforcement of the time limits set for the student presentations.

When I asked, what they had done to alleviate some of the time problems, the teachers answered, "Cut back on journal entries." They went from having the students do two every week: one free writing response and one assigned topic, to having them do one, but that caused some negative reactions from some of the students. "Some people were very upset because we cut out the free one and they had to do the assigned topic. That was the one they liked to do and they want their free topic back. And then there are some kids who can't write on the free topic" (TI2, 12-94).

Several writers/theorists have suggested that technology might play a major role in alternative assessments, but just what that role will be has yet to be determined (Worthen, 1993a; Bruder, 1993). In this study, technology assumed only a minor role in helping the team teachers cope with the time problem. The administration has provided all the team leaders with a computer for their room, and all the teachers have access to a computer in the teacher's lounge as well as in the computer labs, when available. The team teachers put the Senior Seminar grades on the computer which allowed any of them to pull up the grades and add to or change the grades rather than each teacher compiling his/her own grades and then aggregating them at the end of the grading period. In addition, by videotaping all the
senior presentations, the teachers could review them when grading, and could share them with future students and others.

Dividing up responsibilities was another means the teachers used to deal with the lack of time. "Each of us was assigned or given a nine weeks so that person is primarily responsible for that nine weeks. Everybody else aids and abets" (TI2, 12-94, p. 11).

The administrators acknowledged that time is a real problem. The principal reflected on what he could do to help with the implementation of authentic assessment and indicated that is one of the reasons he instituted common planning time for the teams. He has also provided emergency support for the teachers as necessary. When the deadline was approaching for the exhibition, and the teachers had not finished grading all the projects and presentations, Mr. Kent granted them a release day, which freed them from their other classes and enabled them to assess the remaining senior presentations.

The teachers requested an aide or someone to help with the typing or putting the grades in the computer, and some of the more routine jobs. The administrators had indicated that senior team teachers need only ask for whatever they want, but have been unable to grant their request for aides or money for equipment. The principal responded to the teachers' frustrations. "There is no doubt there is a need, but we have many, many other things ahead of that." He offered a possible solution:
What we need to do is get a volunteer organization of paraprofessional type people: moms, dads, grandpas, whoever can, come in and assess with us. How rare is it in your experience that you have a volunteer organization that works with seniors? ... I think this group [the senior team] needs to solicit and recruit some parent volunteers to come in and help them do this kind of stuff, to justify the need for it and also to get the job done. (TI3, 5-95)

Central office has dealt with the time problem by providing planning and reflecting time during inservices, but the proposed solution that was mentioned by all of the administrators was the idea of extending the school year for the teachers. This has been done on a very small scale for the 1993-1994 and the 1994-1995 school years by adding three inservice days after the school year ended in June and has been seen as "very productive time." The administrators would like to see the time increased to at least a month. According to Dr. Jones, "We see that June time as real important, that's been very effective for us.... Dr. Getz feels even a lot more strongly about this than I do, that it's not that we are underpaid, it's that we are underworked, in terms of having a full time job twelve months a year." So his proposal for the board would be for the teachers to work til July 15th and come back August 15th. "I know I'm not going to get from June to July 15th, but on each contract, I'll be negotiating, not negotiating, working together with the teachers' organization to keep adding days on and spend my inservice dollars on those contractual days, so a teacher won't have 184, she'll have 194 the first year or two.
And then by the year 2000, we may be at 210, I hope so, that to me, is the biggest challenge in terms of getting our profession to a true profession" (JI1, 4-95).

Mr. Kent's words were almost an echo of Dr. Jones:

I've always believed that teachers were underemployed, those are Dr. Getz's words not mine. People say you only work nine months a year, well, I believe that teachers need to be engaged in their professional development through the end of June, then everyone takes the month of July off, and start back the middle of August. (KI2, 4-95)

How do the senior team teachers feel about this proposal for solving their lack of time? The senior team teachers had not heard about these plans when asked in the spring of 1995. These teachers already spend a substantial portion of their summer working on school associated work; some of it reimbursed, some of it not. For instance, they had requested additional time to work on the Senior Seminar course after the 1995 final inservice. They were granted the additional time and paid from moneys from a general state grant. However, they indicated in an interview that they probably would have met anyway. For them, teaching is a profession and they spend their time and energies accordingly. The administration's plans would not change these four teachers' lives drastically; they already spend most of the month of June on school related work. However, the teachers' reaction might depend on how much time and on how that extra time was spent.
Another problem that reduces to a time issue has to do with those few teachers who are the risk takers, or as Schlechty (1993) refers to them, the trailblazers. The literature suggests that it is important for those involved in new programs or innovations to share their ideas with others at conferences, and other sites, but the teachers, though wanting to attend conferences and share their ideas, are concerned about their students and the disruption it causes to their classes:

There is so much planning going on, that the building is often filled with substitutes which is killing our kids and killing the discipline and killing... it's a problem. So part of our reluctance is if we go to a conference, then four of us are gone and whatever illnesses and then whatever else is going on and yeah, the continuity. (TI2, 12-94)

Mr. Kent also commented on the teachers' absences because of the training or sharing:

I just got the attendance report of all my teachers and Mr. Stern has been out 18 1/2 days this year, whew. He's team leader, department chair, he's doing leadership, very little illness, he's just been out for this training, well you cannot remodel your house without a good, firm foundation and quite frankly we're going to look at some ways to increase, or let me put it this way, decrease the time teachers are out of class. (KI2, 4-95)

In December, 1994, the teachers discussed how, as the teams worked more closely together, the time available for them to meet with others in their departments was rare. Ms. Dearheart explained:
That's really nice for the teams, but I see a real break down in the departments now, And I think that is detrimental. I think there needs to be some departmental communication. There just isn't time for that. (TI2, 12-94)

This problem has been addressed by the administration and plans made for the departments to spend more time together. Mr. Kent elaborates:

I think one of the most powerful things we've done is we have reorganized 6 through 12 into interdisciplinary teams but also we have maintained our department structure, we have department chairs who are 6 through 12 in the four academic areas plus wellness.... Their focus is on curriculum articulation and coordination 6 through 12.... You have to have departments and teams, it isn't an either/or, it's both. (KI2, 4-95)

All of the problems related to time might take on a different perspective if we consider Frederiksen's (1984) idea in calling for alternative assessments. He suggested that although these situational tests are more time intensive, and costly, if one takes into consideration that they are a part of instruction, then their price doesn't seem so high. This idea is reinforced in the district's Strategic Plan, "The curriculum is inseparable from assessment" (Johnson, 1994, p. I-10).

(2) Interrater Reliability and Validity

After time, the next most debated issue surrounding alternative assessments is that of reliability. Herman and Winters (1994) report that interrater reliability is achievable, but that the
bigger problem remains in the costs of achieving it, that is, "staff training, development of task specifications and prompts, administration of portfolio records and their storage and scoring" (p. 54), as well as "opportunities for professional development, ongoing support, technical expertise and time for teachers to develop, practice, reflect upon, and hone their instructional and assessment expertise" (p. 54).

During the first year, the teachers had seemed unconcerned that there was little consistency between the raters. When asked if they had discussed what they were looking for when they awarded the various scores 1, 2, 3, or 4, they responded:

Ms. Little: Why would you want that? Why would you want it honed down to that?

Researcher: Well so there is some consistency among the raters because that is one of the problems they see in this with subjectivity. Inconsistency of raters makes it unreliable because the students can come back and say, "If he thought I got this and she thought I got that, then gee, I can say I deserve something else."

Mr. Stern: But if you give them an average

Ms. Little: I think that's right.

Mr. Stern: A difference in philosophy, there's always going to be differences in philosophy and whenever, those aren't going to go away, but you'll end up with an average which is probably very justifiable.

Ms. Little: That's right.
Mr. Stern: In other words, someone may grade too easy and someone may grade too hard, but then when you meet in the middle and I try to

Ms. Little: You may not be looking at the same thing

Mr. Bradford: The danger in that was, with me for instance, that I graded the kids higher than my colleagues did, but the advantage of that was, that so did Mr. Stern, so that at least, they were represented by a higher score in each team.

Ms. Little: It's a male thing.

Mr. Stern: It is a male thing.

Mr. Bradford: Well I just, you know, you interpreted the numbers lower

Mr. Stern: Next year going into this, I think we're going to have a clear picture in mind of what we are looking for because we had no clue. We had no clue going in as to what we were going to get. We really didn't know how these presentations were going to look.

Mr. Bradford: Exactly.

Mr. Stern: So next year we will have, we have videotape, we know what they were like this year. We will be able to say, "Well this is a dynamite presentation. Here's some things that you should be looking for." I think we can give some do's and don'ts. Ms. Little's already gotten that, she gave them a beautiful paper which talked about some of the expectations and went on about some of the things that should have been done and weren't done. (TI1, 5-94)

But during the second year, the teachers encountered a problem which caused them to revisit their scoring rubrics. The problem which the teachers called "a problem with
miscommunication" was also a problem with interrater reliability, and was described in Question II, page 190 of the Dissertation in the discussion on the changes made by the teachers. The teachers didn't recognize it as such, but handled it by discussing the inconsistencies among themselves, and increasing the number of points on their scale. They used different numbers of points for the various dimensions (See Appendix E). For example, the project/product was worth 70 points; 10 points for Value of the Product; 50 points for Quality of the Product; and 10 points for the Appropriateness of Project/Product. Herman, Aschbacher, and Winters (1992) suggest that "more scale points enable you to identify small differences between individual students and may provide more diagnostic information than a reduced scale" (p. 66), however, they caution that "it takes longer to arrive at consensus about how to assign scale points when there are more points to consider" (p. 66). They also suggest that when using multidimensional criteria to rate a performance it is helpful to use the same number of points for each criteria or to create "different size scales to reflect relative value or weight" (p. 67).

The scoring rubric created for the judges the second year was also quite different from the previous year's, and enabled the judges to be more consistent and much more efficient. During the first exhibition, the judges were not given any instructions on how to execute the scoring rubric, but for the second exhibition, the teachers refined the rubric. They created a ten point scale for the
judges with "one" being poor, "five" being average, and "ten" indicating excellent, with the total possible points for Project/Product being "thirty" instead of the "seventy" on the teachers' scale. The rubric was sent to the judges ahead of time with instructions that they would mark the scores as the students presented. After the presentations, the score sheets were collected, tallied on the computer, and the winners selected on the basis of the total scores.

Here is a case of the teachers struggling with an issue with which they are unfamiliar, but trying to deal with it in a logical manner. Remember, when the role of assessment was removed from teachers' hands over 60 years ago, teachers were still expected to do the day to day, classroom assessment, often creating teacher-made tests with no instruction in dealing with issues of reliability and validity. The assessment course had dealt very generally with assessment to familiarize the teachers with the concept of "why" we need to change or "add to" the assessments to match "what" and "how" students learn, and very briefly with technical skills. The terms reliability and validity were mentioned in the course, but because of limited time and the scope of the course, the teachers did not have the opportunity to work with these issues and get feedback on how to incorporate these into their own assessments.

In an effort to deal with the issue of responsible and fair assessments, the administrators had the teachers turn in their
semester exams for review. Dr. Getz explained that regardless of whether one is doing authentic or objective testing, he expects good and appropriate assessments:

I want them to be the best semester exams that they can be and I want them to contribute to the cycle of improving instruction not just sorting. (GI3, 4-95)

The principal admits, however "we haven't even dealt with the issues of reliability and validity yet" (KI2, 4-95). Although the push is there for the teachers to look at the tests and analyze them, for item difficulty, reliability and validity, they need an opportunity to learn why that is important and how to do it.

During the first year, the teachers did not recognize a need for consistency, as noted above. The second year, when their intuition and experience told them something wasn't quite right, they tried to make the assessments more reliable, but they still were not interested in checking to see just how reliable the scoring rubrics had been. It would be helpful if the teachers would calculate the interrater reliability among the teachers as well as between the judges, and the consistency across students and across years, to satisfy themselves, students, administrators and the public that this is, indeed, a reliable method of assessment. Or in the event that the reliability coefficient is lower than acceptable, the teachers could make revisions and modifications to the rubrics.

219
Validity is another issue of concern when creating assessments. As mentioned in the literature review, Frederiksen and Collins (1989) have referred to systemic validity and argued that tests should be developed which measure what we value and what we want students to know and "thus reflect and support the development of the aptitudes and traits they are supposed to measure" (p. 28). The teachers of Senior Seminar have tried to create activities and assessments that are both meaningful, useful, and valuable for the students and which assess those areas which they value and want the students to be able to do such as oral and other communication skills.

Linn, Baker, and Dunbar (1991) have addressed the issue of judging the quality of assessments, and suggest that traditional criteria need to be expanded. They propose eight additional criteria: consequences including directness and transparency; fairness; cognitive complexity; content quality; content coverage; meaningfulness; and justification of cost. In the most recent assessments, the team teachers have tried to address or at least consider the criteria of meaningfulness, consequences, and fairness, but not content quality, content coverage or cost.

(3) Dealing with the Unknown

Another issue that these teachers had to deal with was the fact that there were few, if any, models and guidelines for them to follow as they began the new course and planned and
implemented the activities and authentic assessments. Many times during that first year, the teachers mentioned that they were learning along with the students, or they "had no clue," or they were "floundering."

The idea of doing something with seniors to assess what they have learned and what they can do in certain arenas such as writing isn't new. There are many schools which require the seniors to produce portfolios, or theses, or take part in an exit exhibition (McDonald, 1993; Darling-Hammond, Ancess, and Falk, 1995), but the teachers could not find any other school which incorporated all the elements that they envisioned for their Senior Seminar. They read about and even visited other schools and their programs. They took the assessment course and did the follow up independent study. They even felt that they had the support of the administration, but nonetheless, they struggled with creating and designing, from scratch, a whole course and the consequences that go along with that. They worked many extra hours on their own time. Discussing the authentic assessments, where and how they learned the most about them and how to do them, they remarked:

Ms. Dearheart: Trial by fire.

Ms. Little: Yes, probably just doing them.

Mr. Stern: A lot of it is intuitive in terms of, we've seen a lot of things and we've heard about a lot of things. So intuitively, we know the community has to get involved,
that's very important. And I think we saw a lot of it this year with the kids going to different businesses and such where the possibilities blossom out there. It's almost a multiplier effect or something like that, where you start with a small nugget and it kind of opens up. That's the kind of thing you have to experience and see. (T13, 5-95)

(4) Communication and Education of Public
Another problem when introducing authentic assessments into a district or a classroom has to do with communication and educating the students, parents, and the community about what authentic assessment is, as well as, its usefulness and benefits. The administration has attempted to communicate with the public as to what they are planning and trying to do in terms of changes in the district. A booklet entitled *K-12 Program of Studies* (1995) which explains what students are expected to learn at each grade level was printed for the parents and the public (For critique on this, see Question 1, p. 175). The New Albany Company put an ad in the local newspaper referring to the new K-12 program of studies. However, there was no mention of authentic assessment in the ad. In addition, a Public Academy has been created as a vehicle to meet with the public to discuss what is happening in the school district. When the first meeting was held in April, 1995, twenty-five people attended and expressed an interest in having subsequent meetings. Although only six people showed up for the May meeting, additional meetings were planned for fall of
Dr. Getz explained the district's attempt to communicate with the parents:

You still have to be responsive to the customers out there, the client base is still the parents out there and while we are educating and bringing them along if they perceive that you're still not doing some of these things effectively your credibility is shot and you're never going to move them on that continuum. Look at ... some of the difficulties have happened by people trying to take this on too quickly. (GI3, 4-95)

The team teachers met the challenge of talking with the parents about the course and the type of assessments they would be using through a parents' meeting early in the second year. The teachers had heard rumblings of dissatisfaction about the course through some of the student journal entries, and they were not sure what concerns or feelings might be brought up at the meeting. However, nothing negative was mentioned at the meeting and the teachers carefully avoided certain topics.

Students from Senior Seminar performed community service as part of the course which put them out in the community and helped create an awareness of what this class was doing. Likewise, each student was encouraged to have an outside project co-advisor; someone out in the community who was an expert in the area in which they were working. This contact also proved a means of educating the public as to the real world flavor and the interesting work the seniors were doing.
Finally, the exhibition itself was open to the public. The school has encouraged outsiders and critical friends like myself to attend the exhibition. It mainly drew parents and siblings, but the judges represented various facets of the community. Their positive comments revealed their approval and praise of the course and what the students were doing. Additionally, the students' portfolios and interviews were formally evaluated by persons from the U.S. Health Corporation, Fifth Third Bank, and the New Albany Company.

The teachers voiced a concern that had to do with communicating the overall importance and necessity of the course to a few parents and students at the end of the second year.

Mr. Bradford: It's been kind of difficult convincing the kids and their parents that this is a required course, and that there is not a way to make it up in summer school. We have to face some of the problems generated from that. Some of them don't take it real seriously even though it isn't real difficult. We've told parents. We sent them certified letters. (TI3, 5-95)

The school and the teachers have tried to communicate with the public about the changes that are taking place in the district including assessment, but are careful, (1) to refrain from anything that might be too controversial and, (2) to maintain their intentions to continue preparing students for the state mandated and other standardized, objective tests.

However, much of what happens to schools is affected by politicians and how they perceive things. Even though they have
no day to day contact with the schools, and so, have little first
hand knowledge of what students are learning and doing in the
schools, they can help or hinder; approve or withdraw approval
for programs and money. The frustration of how to deal with the
politics of school is demonstrated in the following comment:

I was told that the legislators are upset because teachers are
going places. They are going on these wonderful trips is the
way they expressed it and I guess that really annoys me.
Any other profession, if they want to go participate in a
conference, it's expected, but for some reason teachers aren't
supposed to be able to do that. The other thing is, I'm
hearing things like with this indicator push, they want to see
big changes in student performance, and you're supposed to
be able to start demonstrating this now, that is 1) not
realistic, 2) if you're looking at the kind of information they
want, which is stuff that is very measurable and very
quantifiable, the things you can measure may not be the
most important changes that are happening. We've got
some really complex things happening. How do I, in a
measurable form, explain to a legislator the results of the
Senior Seminar group? (GI3, 4-95)

(5) Support

Support from the administration is one of the most
important factors contributing to the success or failure of
programs. Darling-Hammond and McLaughlin (1995) see the role
of administration as being supportive by "rethinking schedules,
staffing patterns, and grouping arrangements to create blocks of
time for teachers to work and learn together" (p. 601). The New
Albany school administrators have been extremely supportive of
the senior team teachers. They have paid for attendance at
225
workshops and the assessment course through the fee waiver agreement with Ohio State. They have encouraged the teachers to try new things, have given them planning time by rescheduling and providing time during inservices. They have praised the teachers for their efforts and allowed them time to attend conferences and visit other schools to share their ideas and their work.

They have encouraged the teachers to take risks and be innovative. According to the teachers:

Mr. Stern: I think they encourage, they don't say "You can't do something." ... They've always supported. They say, "that's great. That's a good thing. Keep going with it or something like that."

Ms. Dearheart: I think even if they don't necessarily agree, they say go ahead and try it, and then evaluate it. I don't know that we've ever been told, "Don't do that."

Mr. Bradford: We've been given a lot of room to experiment.

Ms. Little: Carte blanche. (TI2, 12-94)

Dr. Getz offered several ways in which the administration was willing to support and help the teachers as they learned about and implemented authentic assessments into their classrooms. Providing a framework which supports the needs of the teachers; examples and models for use by the teachers; as well as time for sharing are some possibilities:

We want to generate a framework that will support the evolution... We obviously want to look at how we can
address modeling in a way that is economically feasible and cost effective and all of those things. I think the technology has some promise here. I think that videotaping people actually doing authentic assessments and creating a situation where people can share those videotapes is important. So I think once they see one model, they will be able to generate additional ones. So trying to establish within this framework, ways of communicating examples and things that work and things that don't work is very important. ... One way of sharing the collective vision is to see how many more people we can get on board. (GI1, p. 4, 2-94)

Administration has also supported the teachers as they struggle with time for reflecting and planning as discussed in Question I. The long term plan of central office is to increase the number of paid days for all teachers so that they may be able to plan, share, and reflect on their teaching practices:

Right now every time we do those three days at the end of the year we have to find money to support it. We hope that during the next round of negotiations we can build that in as regular contractual days so that we're not playing this 'find the money'. And also so we can reimburse the teachers at a rate that is commensurate with what you are asking them to do. So we'd like to be able to build those into the contract as a regular part of the day. The other thing that we're doing is our inservice days are now primarily sharing times you know rather than being a delivery vehicle for some snapshot of an item, we want their work plans to generate their own approach to staff development and then our inservice days are going to be more about communication, sharing or revisiting the vision rather than training days. And we always provide time for teamwork. (GI3, 4-95)

At the building level, the high school principal created the teams and has worked with the class schedules to create
planning time for the teachers (individual, team, and departmental) for next year. He has also provided release time when they needed it to finish grading all the projects and presentations.

Only in a few instances did the administration falter in the area of support. Early on, the administrators talked about the Sherwood School district and what was being done there, creating anxiety and frustration among the New Albany teachers. In addition, neither of the central office administrators were able to attend the first exhibition and the team teachers were noticeably disappointed. The messages the teachers get from things administrators both do and say are very strong, but the "action" messages seem to have the most impact on the teachers' attitudes and feelings. Although the administrators had told the team they had only to ask for what they wanted, when the teachers requested an aide, they were told that there wasn't enough money. Another example of this was observed when several teachers mentioned that the administrators and the professor of the assessment course didn't give them enough credit for what they knew and when the teachers indicated that they felt that the administrators thought that New Albany teachers were not as good as the teachers at Sherwood. Whether any such comments were actually made by the administrators or not, nonetheless, these were the perceptions as seen by the teachers based on what was being taught and the changes the teachers saw being made.
QUESTION 4: What are the strengths, weaknesses, and constraints entailed in alternative assessments, according to teachers, administrators, and students?

The methods and beliefs of what would be used in the district in terms of assessment (listed in the Strategic Plan and on page 95 of the Dissertation) portray assessment as something ongoing, using multiple techniques, being personalized, being engaging, interesting and challenging, and "the limited role of established standardized tests" (Johnson, 1994, p. 1-10).

Teachers

The senior team teachers in this study do not see authentic assessment as the answer to all the problems, but as a way of capturing a more comprehensive picture of what students have learned and can do. They feel that objective, paper and pencil measures still have a place in their repertoire of tools for assessment:

Mr. Stern: I used them [authentic assessments] before you know. I thought they were authentic, like in the mock trials and things like that and I don't see anything wrong with them; I think they have a place and time. I think you still need the more traditional assessments, too. You just can't do just authentic. I don't think; I think what it does is gives it balance; it gives more interesting kinds of things for the students to be engaged in, other ways too; it does both. I don't see it as an end all and be all, but I do see it as an important tool.
Ms. Little: I think that was a nice word you used, the balance

Mr. Stern: It does; because you do a lot of that stuff

Ms. Little: Yes, but, you know, I still do; I still think you can still teach with tests and quizzes, I think that is still a viable teaching method too, but it used to be all or nothing, but I do agree you can't swing the other way; don't let the pendulum go totally to authentic because that's not good either; and the colleges aren't doing that anyhow, so

Mr. Stern: Well maybe

Ms. Dearheart: I agree, in my classroom I need to test skills and that's the best way to do that is tests and quizzes; and then the other assessments come from experimenting

Ms. Little: or applications.

Ms. Little: Obviously one of the strengths (of authentic assessment) is that the things are usable. The skills that they are picking up are usable. We passed out portfolios yesterday when we had our team award assembly ... and one of the girls had gotten hers earlier because she was going on a job interview at Bank One and needed to have it completed. So they are using the things they are creating.

Another strength the teachers noted about authentic assessments is that they allow students to work within areas which are meaningful to them and in ways that feel comfortable to them. Additionally, one teacher commented:

Ms. Dearheart: I think some kids identified strengths they didn't know they had during that whole project and the
assessment. I'm thinking of one student, I'm not sure that he realized that he had the potential to do a process journal as well as he did. (TI1, 5-94)

The teachers also recognize that authentic assessments are time consuming in terms of both planning and grading. Mr. Bradford commented:

On the down side, we have found it difficult to have the students do the projects over if they were substandard. There just wasn't time, and we may have to come to some sort of compromise there in doing that. Also, if we were to fail a student in Senior Seminar and they would have to do something in summer school to make up for it. We haven't dealt with that yet either, hopefully we won't have to this year, but there is no summer school mechanics to make up.

The teachers believe that a balance of various assessment techniques is the best way to capture what their students are learning. They see strengths and weaknesses in both the more traditional objective tests and the authentic tests, but feel by using a combination and determining which is more appropriate for each particular situation, they have the best of both worlds. However, they do not understand the administration's push for item analysis, as illustrated in an interview with the senior team teachers:

Mr. Stern: They do want exemplars of things. They want us to document and keep exemplars.

Ms. Dearheart: That's become the major push.
Ms. Little: Item analysis which is very traditional.

Mr. Stern: It doesn't fit well with authentic assessment.

Ms. Little: No, it sure doesn't, but I'll bet for our semester exams we will be asked to do that. (TI 2, 12-94)

Administrators

The administrators, likewise, see a need for a variety of assessment techniques and feel that what is needed is good, sound assessment. Dr. Getz offers the justification for doing item analysis:

Wherever on the continuum I am generating an assessment I want it to be a good assessment. I wouldn't say that this fits into the move toward authentic, this just fits into trying to make all kinds of assessment that we do, be good and appropriate assessment, so given the reality of a world that still expects you to give semester exams, I want them to be the best semester exams that they can be and I want them to contribute to the cycle of improving instruction not just sorting. So by going through the item analysis that is related to the semester exam it's trying to improve that instrument and provide the vehicle by which the instrument can have some impact on instruction.... It is an effort also to get people to understand that multiple choice exams are not objective, scoring multiple-choice exams is objective, creating multiple-choice exams is extremely subjective, and the item analysis is another way of altering your perspective to see that, gee a whole lot of students missed this item, now how important was it really? (GI3, 4-95)

Mr. Kent, the high school principal concurs that there needs to be a balance of techniques:
People are beginning to understand that it isn't an either or, it's both, it's both traditional paper and pencil and authentic assessment component. (KI2, 4-95)

The superintendent also stresses this point:

It's the balance. It's still, there are multiple assessments and I don't believe that authentic assessment ought to be the main linchpin. It needs to complement and balance the graded course of study objectives, of which many are, frankly, recall and demonstrating mastery of certain facts. (JI1, 4-95)

Students

The students in the first year Senior Seminar course whom I interviewed made these comments about the projects they had done and what they thought about the various kinds of assessment techniques:

Student 1: I think that the majority of the people would have rather of taken the exam

Student 2: Plus we had all our other work and the exams just involve what you are doing, and so you take it at the end, so I think a lot of people would rather have the exam

Researcher: So you think you might rather have taken an exam because it is easier.

Student 3: It's easier.

Student 4: And it's the information you have already studied.
Student 2: And you can easily take two days off work to study or two days out of your extra curricular activity to study for an exam where this you can't

Researcher: So in which respect do you think you might learn more?

Student 1: We'd learn more I'd say, this.

Student 2: You learn more about one thing.

Student 3: Even though we didn't do that much on the project, you know, we'll remember that little bit that we may have learned there, more than if we had taken a whole entire test.

Student 1: That's true

Student 5: I think I might like a test or multiple choice test or something better, I mean after I was done with my project I was happy and, I was glad that I got it done, but it just takes a long time, it takes a lot of time and patience and it just, I mean hard work really, but I think you learn more from a project than you do a test, I think you learn more. (SI1, 5-94)

The 2nd prize winner who built the car indicated that he preferred an assessment like the project: "Yeah, cause I can do just about anything but take a test. It's about what other people have learned, not what I've learned." The project "just combined my hobby, and my schooling, and what I'm probably going to do for the rest of my life, all together" (SIA, 5-94).

Researcher: "Has this helped prepare you with your future plans and how?"
Student 5: Yes, because if I want to be an English major it has given me experience on writing and stuff and if that's what I want to do for the rest of my life then at least I'll have something to look back on and know that I did something to help me. I mean before this I really never wrote that much and I never wrote all the essays and poems that I did, so I think it helped me on how to write and how it was wrong and how it was right and stuff.

1st place winner: I really like the way I did [the project], I really learned more because I was responsible for more; there was more that I had to do, whereas a test they just measure what you learned, really there is nothing that can measure what I learned from this project because I learned a substantial amount both about my community, about what I want to pursue in life, I mean there is a lot that I learned from it.

The students who received good grades on their projects and presentations were more inclined to feel positive about authentic assessments and feel that they learned more by using this means of assessment rather than objective tests. The students who did not do as well on their projects were inclined to feel that this type of assessment was less fair, took more time, and were much harder. Their reaction was partially a result of how these assessments were introduced to the students and the fact that many of the requirements for using them were not in place early in the year. For instance, the students did not know the criteria and the standards in advance nor did they have exemplars.
Student 1: I think what they expected from our projects were different from what we expected. What we thought they expected, we did, but they expected more than what we had thought...

Student 2: I think there was too much of a wide range of projects. Cause some people like built a car, and some people made a pamphlet, and you can't really compare.

Student 1: Now, if they had said, "You have to do an extraordinary project," then I'm sure we would have, but.

Student 3: They gave us the grading system a month before we had to present it, not even.

Student 1: And we had this whole year, and they gave us exactly what they were looking for like two weeks before.

Student 2: And it was too late to change our whole project. They didn't even give it to us when we had the project, they waited til... (SI1, 5-94)

This lack of clarification was due to the fact that this was the first year for the course. The teachers did not know what they wanted. They were learning, too. It was during the assessment course that the teachers thought about the criteria and the scoring rubrics and presented these to the students. The students admitted, however, that they probably learned more, at least about one thing, through these authentic assessments.

All of the second year students who were interviewed agreed that the assessments were fair. Although two of them felt that the criteria were initially unclear, with feedback on their work from the teachers, they understood what was expected. The
students in the second year expressed similar attitudes toward objective and authentic assessments as the first year students. Mary responded, "I like objective because they're easier, but the projects are geared more towards what you are going into" (SI2M, 5-95). Phyllis, one of the finalists, added, "I don't take tests well, I learn the material and it's all in my mind, but when I go to take a test I get too tense and too nervous.... With my project, I could relate to it, and the problems within that incorporated a lot of different skills, but I could solve them because it was something that dealt specifically with me.... It was definitely a learning experience" (SI2P, 5-95). Joe, a student who received an honorable mention remarked, "It lets you pick what you're interested in and lets you show what you can shine on instead of filling out the bubbles A, B, C and it gives you a chance from the inside to make a product that caters to your needs" (SI2J, 5-95). It is interesting to note that this comment almost paraphrases Sizer's assertion that tests should permit students to "show off what they are able to do, not be a trial by question" (Sizer, 1984, p. 64).

One problematical issue which was mentioned by several of the students had to do with money. Many of the projects involved thousands of dollars and this concerned and worried the students as to what kind of projects were expected and how students could raise the money needed for such projects. For instance, the first
prize winner had spent several thousand dollars remodeling the basement and the student who rebuilt the car also spent several thousand dollars. This seems to be an equity problem and needs to be addressed.

The perceptions of the three groups: teachers, students, and administrators on this question are similar. They see the benefits and the need for authentic assessments as a part of instruction and a way to allow students to become active learners engaged in their own education. The teachers and administrators feel that a balance of assessment techniques is necessary to capture a complete picture, that each has strengths and weaknesses and can be used in the school setting to complement one another. This seems to be consistent with the literature. Stiggins (1986, 1991, 1994, 1995) has argued for sound assessment using a variety of assessment tools for the appropriate purposes and in the appropriate situations. Likewise, Wiggins (1993a) believes in sound assessment using a variety of techniques which respect the rights of the students. Haney and Madaus (1989) urge that, "more than new forms of assessment, what is needed is a refusal to accept bondage to any single technology - no matter how useful it may be in a particular instance - and thoughtful selection of different kinds and mixes of assessments for different purposes" (p. 687).
Conflicting Perceptions

Throughout Chapters 4 and 5, I have pointed out the differing perceptions and viewpoints of the participants in this study. Now, I bring these together to illustrate how they cross the boundaries of the various positions within the community and school. These perceptions include:

- The different perceptions between oldtimers and newcomers.
- The different perceptions between the school board and superintendent and the conflict that it created for Dr. Jones.
- The conflict between administrators and teachers on when the course would be offered.
- Different expectations and perceptions of administrators, professor, and teachers concerning the assessment course.
- Different expectations for projects between teachers and students, and the different perceptions of which journal entry should be cut.
- Perceptions of administrators and teachers as to what is important in reference to the Exhibition.
- Perceptions of what constitutes good assessment and how you get it.
- Different perceptions on the part of students who got good grades on projects and those who did not.

To make some sense of the differing perceptions and conflicts between the participants in this study, the ways they were handled, and how some came to be resolved, I turn to
constructivism. Having entered this study from the epistemological standpoint of a constructivist, I ask myself, "If students have different experiences and construct their own knowledge, how can we rely on objective tests to find out what students know? I believe that authentic assessment offers a plausible alternative. Authentic assessment is consistent with a constructivist view of knowledge and theory of learning. By using assessments which are meaningful, are used in context, allow students to choose alternative means, and encourage collaboration and cooperation, teachers give students the opportunity to construct and negotiate meaning for themselves.

When I began the research in the fall of 1993, I had no clue that constructivism would also provide me the means for understanding and explaining the conflicts and different perceptions I encountered during this study. I was continually asking myself, "Given this situation, how could it be perceived so differently based upon one's role or position?" As I began to analyze my data and had the opportunity to get some distance from the research site, I began to see patterns.

Objectivists and some constructivists acknowledge the existence of "reality." However, these constructivists, while not refuting the existence of reality, suggest that each individual's reality is somewhat different because it is based upon his or her perception of that reality. Teachers and administrators, as well as students, construct meaning for themselves based upon their prior
experiences and prior knowledge and within a particular context. Certainly, a teacher's experiences and an administrator's experiences are different.

Furthermore, the contexts within which they operate are substantially different; the teacher's primary concern and day to day reality has to do with the students within his or her classroom, while an administrator's day to day reality encompasses a broader focus including many classrooms, students, teachers, parents, other administrators, board members, and other community members, as well as political realities. Duffy and Jonassen (1992) further suggest that, "There are many ways to structure the world, and there are many meanings or perspectives for any event or concept" (p. 3). Is it any wonder that in a given situation, the perceptions of teachers and administrators might start out differently?

So how can total chaos be avoided, and how does one come to some resolution that isn't in total disagreement with the rest of the world? According to Bednar, Cunningham, Duffy, and Perry (1992), it is through negotiation and collaboration with others that we are able to arrive at some consensus. They suggest that an important component of learning and understanding is the ability to evaluate alternative perspectives looking at their strengths and weaknesses and then choosing a perspective "that is most useful, meaningful, or relevant ... in a particular context" (p. 28); this is done through social negotiation.
The challenge for the teachers and administrators in this study became one of negotiating, evaluating the different viewpoints, and collaborating toward some useful perspective that would work for their common interest: the students. During the second year, this is what happened in most cases. The teachers and the administrators were seeing eye to eye on many issues and were working together rather than each group working on similar things but separately. There was more interaction, more communication, more listening to and evaluating of others' ideas as they were working toward mutual solutions for the problems they encountered while implementing authentic assessments.

SUMMARY

Although school reform has been a national issue since the 1980s, New Albany's journey began with the development of a new suburban community around an established, rural community and the idea of a new world class school system to serve this proposed larger community. The New Albany school district's administrators and teachers have progressed significantly on their journey from a traditional model of school to a model based on a constructivist view of learning which includes the use of authentic assessment. However, they are still in a transitional stage and have several areas and issues which remain to be addressed and/or completed.
Administrators

The superintendent who was hired at New Albany possessed knowledge and experience in the use of current theories and innovations in education. He was very aware of the latest research on learning and how this related to teaching and assessment. He created a very detailed strategic plan which would put his knowledge and beliefs into action. The director of curriculum and the high school principal were also very knowledgeable concerning the most recent innovations and theories in education. The changes planned for the district were "top down," originating from the school board and the superintendent.

Initially, many changes were being implemented in the district without teacher input, and this created some problems and resistance among the teachers. Most of the decisions originated in central office. However, communication and interaction between the teachers and administrators have increased over time, with more listening to and evaluating of each others' ideas.

Administrative support for teachers during the introduction and implementation of the changes included scheduling for common planning time, inservice time for planning and reflection, the graduate assessment course, and eventual praise. At the completion of the study, the administrators seemed to be content with the way authentic assessment was being implemented within
the district. According to the strategic plan, this process should be completed by 1996-1997. Dr. Jones noted that some teachers would do it the first year, but he expected that for others it would take five years. However, as the 1995-1997 school year ended, the senior team teachers who are the trailblazers of the district were still learning and experimenting with these assessments themselves.

A change from a traditional system to a system consistent with a constructivist paradigm must include restructuring of the whole system. In terms of day to day schooling, the New Albany school system's structure is much the same as it has been in the past. School primarily takes place within the school building and within classrooms. This is beginning to change for the seniors as they are being allowed to pursue their projects outside of school, but only during the Senior Seminar period. Additionally, the concept of an expanded learning community has been at the core of the strategic plan and many of the proposed programs are at various stages of implementation (for example, the wetlands-woodland, environmental stewardship program and the wellness curriculum).

The district's reporting system is still the same; grades are reported in terms of A, B, C, D, and F. However, student led conferences have been expanded into the high school and into the elementary school, and administrators are exploring other possibilities. They have sent teachers to learn about "the
ramifications of using authentic assessment for reporting to the home" from CLASS (Center on Learning, Assessment, and School Structure (GI3, 4-95).

The standardized tests are still in place, and these present several questions and issues for the district. Can teachers incorporate new activities and tasks while preparing students for the many standardized, objective tests? What does the state’s proficiency test mean for the district and how far can they move with it in place? What about the district's hiring policies in terms of assessment literacy?

Other issues have to do with standards which have yet to be firmly established. How will this happen? Who will be involved? The administrators have indicated that communication with parents and the public is one of their prime goals, but thus far, assessment or authentic assessment has not been among the topics of discussions or communications with the public.

Teachers

The senior team teachers have engaged the senior students in authentic activities built around team building, careers, community service, and their final projects. Many of these activities either sent the students out into the community or brought members of the community into the school.

The senior team teachers have begun to clarify for themselves and their students what it is they expect and value.
They shared these expectations with the students and the parents. The expectations and findings from the exhibition were also being shared with teachers K-12, and are being incorporated into the curricula. Although the senior team teachers have developed understanding and more sophistication regarding criteria and scoring rubrics, they need to include students in the development of these. Also, the rubrics should be more specific, especially those for the exhibition judges.

The concept of "redos" was initiated by the teachers to allow students to practice and improve. This is consistent with the philosophy that all students can learn, and it also extends the responsibility for learning to the students.

During the implementation of the authentic assessments, the teachers ran into several problems like lack of time and trying to work around some of the more traditional expectations of the school and parents. To deal with the problems, the teachers began to retreat from some of their original plans that were consistent with authentic assessment. They cut back on redos. They made decisions about journal entries which cut back on choices and decisions given to the students. They avoided involving the students in the development of the criteria because of the time it would take. They built a component of the Senior Proficiency Test into the Senior Seminar exam grade to motivate the students to try harder on the test.
The senior team teachers may be ahead of many of the other teachers at New Albany in their use of authentic assessments as well as in other innovations. They are the trailblazers, the school's "thoroughbreds." According to Worthen (1993b), it is helpful for a school moving toward alternative assessments to have "someone nearby who 'has been there' and can demonstrate new methods to neophytes or critique embryonic efforts" (p. 456), and certainly the senior team teachers can fill that role. They have "been there" and have begun to deal with many of the dilemmas that may be encountered when implementing authentic assessments. However, the senior team teachers are still very much in transition, as is the whole district, and they have several hurdles yet to jump before using authentic assessment to its full potential for themselves, the students, the parents, and the district. They are perhaps at the intermediate to advanced level in some ways, but still at a novice level in others. The main hurdles that remain for the teachers have to do with the scoring rubrics and reliability, self-assessment, and communication. Will they continue to move forward with authentic assessment and take the next steps or will they be forced to revert back to more traditional means? The areas which will need to be addressed in the district to enable the teachers to proceed forward have to do with scheduling, grading, reporting, communication, and informal assistance.
Students

The first year Senior Seminar students and the second year students have had very different experiences with authentic assessments. The first year students did not have clear expectations of what they were to do. This created a situation in which many students experienced frustration and resentment throughout the year. They did the kind of work they had done in previous years and in other classes, and this was not judged to be good enough.

For those students in the second year who had the full exposure to authentic assessments throughout the course, the experience seems to have been more positive. They knew what was expected of them through exemplars from the previous year and through rubrics. They chose meaningful projects and felt that they learned from them. They constructed new knowledge based on their previous knowledge and schooling, their interpretations of that knowledge, and new experiences. Their senior projects and presentations were good examples of authentic assessment; they were meaningful, authentic, challenging, enabling, required collaboration with peers and/or co-advisors, and were contextually situated in the real-world.

Consistent with what the literature has shown about authentic assessment and its affect on learning were the comments from the students from both years (Bennett, 1993). Students indicated that multiple-choice tests were easier, but that
they probably learned more from the authentic assessments. The literature also has suggested that assessment has an impact on students' self image (Collins, 1990). Many of the students expressed positive feelings about the work that they had accomplished. However, the students did not seem to understand the purpose or importance of self-assessing their work.

Overall, the administrators, teachers, and students in this study see the benefits and importance of including authentic assessments into their repertoire of assessment techniques to improve student performance. Authentic assessment has been only a part of the district's strategic plan to bring the school in alignment with a constructivist paradigm of learning. The process of moving from a traditional model of school to a constructivist model requires major structural changes in the school system. Some of these changes have been made, some have been considered, and others have been cautiously avoided for the time being. And, all of this is consistent with a constructivist view of the world.
CHAPTER 6
CONCLUSIONS AND RECOMMENDATIONS

The findings of this study lead to the conclusion that even when the administrators of a district moving toward a paradigm shift are knowledgeable about professional development and the latest theories of learning and assessment, and have created specific plans to bring about that shift, the journey is not necessarily smooth and problem free. This case study describes the introduction of authentic assessment into a school district and the implementation of authentic assessment into a particular class, the Senior Seminar. The data describe the problems, attitudes, struggles, and issues which were encountered by teachers and administrators during this process. As pointed out by Fullan and Miles (1992), "change is a journey, not a blueprint" (p. 749). Just as we cannot prescribe exactly how to teach and how to learn (Streibel, 1989), neither can we prescribe how to make changes. Each situation is different having its own unique setting, unique set of circumstances, and players with their own unique experiences and knowledge.

Likewise, there is no blueprint for introducing and implementing authentic assessment. There are suggestions from
those who have tried it and those who have watched others try it, and there are descriptions or reconstructions of the paths taken by others. I offer this case study as one description of a particular district's journey, and leave it to the reader to take and use whatever might be helpful in preparation for a similar journey, whether it be a journey involving a single classroom, a building, a school, or a whole district.

The findings of this study suggest that authentic assessment is seen by teachers and administrators as a useful and beneficial tool which facilitates student learning and serves as a means of demonstrating the process and products of learning. However, authentic assessment is not without its problems, issues, and concerns. The conclusions of this study are similar to what is being reported from other studies on authentic assessment. Time, support, communication, reliability, philosophical understandings, and the different perceptions and beliefs of participants in the various roles are the primary issues and concerns documented. Constructivism provides a means to understand some of these problems and issues, and provides the theoretical base for the move to authentic assessment. If administrators and teachers truly embrace a constructivist paradigm, their roles should be less authoritarian and more coachlike: offering suggestions and guidelines, empowering teachers and students respectively to make decisions and take more responsibility. In this study, administrators and teachers developed a more facilitating role,
but there were still instances in which they maintained and exercised their power.

Teachers and Time

A shortage of time is at the forefront of problems with which the teachers contend. Consistent with the time issue mentioned in other studies, the amount of time necessary to plan, develop, use, and grade authentic assessments have frustrated the teachers in this study as well. The senior team teachers have dealt with time problems by cutting back on some of their initial requirements and expectations such as number of journal entries and student opportunities to redo work. They have divided up responsibilities with each teacher assuming the leadership role for a nine week period. They have used technology in a limited way by employing the computer to keep track of all four teachers' grades for the students, thus making calculations of final grades more efficient.

There are several other means which the teachers should consider to deal with the time dilemma. First of all, the teachers in this study have not used student self-assessment and peer-assessment sufficiently. Many of their students have not understood the purpose of assessing their own work and have been unsure of how to do it; instead, they have relied on the teachers for all their critiques and feedback. If self-assessment and peer-assessment were appropriately and regularly practiced,
and not left as something done when the work is completed, the teachers could be spared an enormous amount of "regrading" as the students assessed their own and their peers' efforts toward achieving the criteria and standards.

An additional benefit of self-assessment would accrue to the students. It would provide them with the skills and "habit of mind" to assess their own work. According to Wiggins (1993c), "If we want people to gain control of important habits of standards and habits of mind, then they have to know, perhaps first of all, how to accurately view those things and to apply criteria to their own work; they cannot always be dependent upon another person for assessment" (p. 54).

Another way to deal with the time issue has to do with borrowing from others. There are many schools which have adopted various forms of authentic assessments. So, even though other schools may not be doing all the activities of the Senior Seminar, the senior team teachers could borrow exemplars, tasks, and scoring rubrics from various programs which are similar, and then adapt them into their own assessments. Dr. Getz had proposed a framework which would provide a means for acquiring models and exemplars through networking and to begin archiving them. This, however, is still in an early stage.

The teachers need to understand the advantage of developing or at least adapting their tasks, criteria, and rubrics to the needs of their students and themselves. Although having
ready-made models to use may reduce the time required somewhat, it may also displace one of the benefits of using authentic assessments. It has been argued that at the heart of the issue of alternative assessments are the questions: What is it we value and what do we want students to know and be able to do? How will we know when they have done it? What kind of performance will suffice? What does "good" look like? What kind of feedback can improve student performance? The process of thinking about these questions and the kinds of tasks the students might do, and then creating the performance criteria, allows teachers to struggle with, and "articulate and come to agreement on what they value" (Arter, 1993).

It is vital to keep in mind that authentic assessments are part of the instruction-learning process, not something to be tacked onto the end for the sole purpose of determining a grade. Nonetheless, additional time will be required for these authentic assessments. Those who are planning to implement authentic assessment should be aware of the additional time required, and start small, perhaps adding one authentic assessment at a time.

Administrative Support for Dealing with Lack of Time

The administrators in this study have helped teachers deal with the time problem by: providing inservice time for reflection and planning, rescheduling classes to provide common planning time for teams and individual planning time, and providing
emergency release time. Although support in terms of praise and encouragement has been forthcoming from administrators, additional help in terms of human power and money also would be greatly appreciated by the teachers. The teachers' request for an aide was denied because of lack of funds.

Shepard (1995) has reported that it is following their initial trials and experimentation that teachers may have questions or need further clarification from experts. Therefore, informal assistance is recommended to deal with these questions and problems encountered by the teachers as they struggle to implement the assessments. A resource person analogous to a curriculum director would be useful for this purpose.

Recommendations

There are a few suggestions which might provide additional opportunities for teachers to plan and grade. Although the administrators have worked with the schedules to allow for common team and department planning, they might consider block scheduling. Various forms of block scheduling such as the 4x4 plan, trimesters, and designating one day of the week for in depth projects have been tried elsewhere with success (Canady and Rettig, 1995; Stumpf, 1995). Many suggest that these block schedules allow for more innovation. Teachers have reported such benefits as: allows for more hands-on projects and more in depth study, encourages cooperative learning, and allows for more
student involvement and personalized instruction (Canady and Rettig, 1995; Hanson, 1995; O'Neil, 1995). Block scheduling can also allow for more flexibility because teachers have fewer students which means more unassigned time to use for planning, grading, and collaboration with other teachers (Stumpf, 1995).

Extending the school year commensurate with pay adjustments may also be a viable alternative, as New Albany administrators have already suggested. The extra time in the summer gives teachers the opportunity to attend relevant conferences and workshops; to plan, discuss, share, and collaborate with other teachers within and across grades and disciplines; and, to reflect on what they have done, what worked, what did not, and what they should do next. This suggestion, however, may be met with resistance from some teachers.

The issue of standards and standardization should be revisited by the teachers in relation to the journal and journal entries as discussed in the last chapter, and in setting the standards for the various assessments. Consideration needs to be given to whose standards will be used: national, state or local. Setting standards should be a joint venture involving teachers, administrators, parents, students, and community members. Once established, these standards then need to be shared with all those with vested interests in the schools. Addressing the issue of diversity and standards, Wiggins (1991) explains, "A 'standard' is an exemplary performance serving as a benchmark" (p. 19).
"Excellence is not a mere uniform correctness but the ability to unite personal style with mastery of a subject in a product or performance of one's design" (p. 19).

During the study, the senior team teachers developed greater expertise in developing the criteria and rubrics as well as scoring the assessments. However, they are still in a transitional stage and need to continue to learn and expand their sophistication in this area. They have progressed from where they were, but in terms of establishing standards and producing reliable and valid assessments, they need to move further.

It would be helpful if the teachers would calculate the interrater reliability among the teachers as well as between the judges, and the consistency across students and across years, to satisfy themselves, students, administrators, and the public, that this is indeed, a reliable method of assessment. Or, in the event that the reliability coefficient is lower than acceptable, they could make more revisions and modifications to the rubrics. The reliability of the rubrics would probably be increased by using a numerical-qualitative scale that was more descriptive; specifying what would constitute a score of "one" and what a project that warranted a "nine" or "ten" would look like. The teachers may have discussed this among themselves and come to some agreement, but the judges have no way of knowing what constitutes (in the eyes of the teachers) a poor, average, or excellent project or product.
It is critical to ascertain the reliability and validity of these assessments, especially, if they are considered high stakes. Since the Senior Seminar course is a requirement for graduation, the authentic assessments used in this course, and particularly, the project and presentation, fall into this high stakes category. The administrators need to provide the teachers with further assistance in the technical aspects of assessment, either through inservice sessions or by providing an individual to help teachers on a one-to-one basis as needed or some other means.

If parents and the public are to trust and accept these new assessments, they must also be provided with adequate reassurances and clear understandings of what the information they receive about their students means. Does it convey the progress made by students or their effort? Does it compare the student's work to other students' work or to a set standard? All of this needs to be clarified for the users of the information.

An exit exhibition has many benefits. It provides a target for students K-12, and it provides a public forum in which students can demonstrate what they have learned. Although the monetary rewards for the best projects, presentations, and journals may have provided incentives for some students at New Albany to do exemplary work, requiring all students to present their projects at an exhibition and before a general audience might provide more motivation for students. Extending this challenging learning experience to all seniors would demonstrate
the expectation of excellence and open the door for all parents to see what authentic assessment is, how rigorous it is, and what students are really capable of doing.

The assessments used in the Senior Seminar are a means to an end as discussed by Darling-Hammond, Ancess, and Falk (1995); but they could go much further. They should "serve as catalysts to bring together teacher and student, student and student, teacher and parent, student and parent, teacher and teacher around examinations of student work and questions of goals and standards" (p. 18). Shepard and Bliem (1995) described how they invited parents into the schools, asked them questions about assessments and reporting, then showed them examples of standardized tests and performance assessments. When presented with both types of assessments, parents showed approval of both types, "giving stronger approval ratings to performance assessments" (p. 31), and indicating that the use of performance assessments "make students think" and "are likely to give teachers better insights about what children are understanding" (p. 31).

The New Albany administrators need to take a more aggressive approach to initiate communication with the parents and the community regarding their paradigm shift and the role assessment plays in that shift including authentic assessment. They have held open meetings, but these have been poorly attended. The superintendent explained that this was his
personal rubric "to get community support." But discussions with the students in Senior Seminar regarding their parents' attitudes about the course indicate that there are misunderstandings, and that the parents need more information and explanations. More communication is the key with more openness and less avoidance of difficult or controversial issues. Communication with the public concerning the changes in the district needs to expand to include discussions about assessment, particularly authentic assessment and its implications for learning.

**Changing Roles**

Moving to authentic assessment necessitates many changes including adjustments in the roles of teachers, administrators, students, and parents. The administrators' roles have changed from the beginning of this study to its conclusion two years later. Administrators have acknowledged the need to step back and provide support for the teachers, while allowing the teachers to become self-learners in relation to professional development. Dr. Jones indicated his role similarly. He explained that his job as superintendent is not much different philosophically than that of a teacher. "My goal is to get to the end of the year, and if I've done my job, I remain silent while the child explains and does and demonstrates and accomplishes those things while the teacher remains silent" (JI1, 4-95).
Teachers moving from a traditional model of school to a constructivist model must change from information givers to coaches and facilitators. In many cases, the senior team teachers have assumed these new roles. They discuss ideas regarding the students' projects with the students and offer suggestions. Additionally, the teachers of Senior Seminar have given up their roles as experts and students choose co-advisors from the community to assist and advise them with their projects.

The role of the community has thus changed in New Albany. However, it remains to be seen how many will take advantage of this new role. The administrators have committed themselves to creating a community of learners in which members of the broader community share the responsibility for educating students and also assume the role of life-long learners which they model for students. Members of the community serve as judges for the exhibition, and the whole community is invited to serve as the audience for the students' presentations. As just mentioned, community members also serve as co-advisors for student projects.

Student roles have also changed somewhat. Seniors are taking on more responsibility for their own learning as they choose projects and go out into the community to gather information and work with co-advisors. They have also assumed the role of active members of the community through community
service. However, they still do not understand nor actively practice the role of self-assessor of their own work.

Parents are crucial to the success of authentic assessment. If they do not understand what it is and why it is being used, and are more concerned with traditional measures and how their district compares with other districts, then this movement will fail. A move to a constructivist model should encourage and include more parental involvement in the learning process. To this point, the parental relationship to schools at New Albany has changed very little.

**Positive Outcomes**

The senior team teachers are continually identifying and articulating what it is they value and what the students should be able to do. An important outcome of implementing these authentic assessments in Senior Seminar has been a change in instructional practices; teachers are sharing expectations with the students and helping them to know and understand what is expected of them through the criteria and scoring rubrics. This is fulfilling the important role of assessment alluded to by Stiggins and Conklin (1992), "It is teachers' classroom assessments that students rely on to help them set expectations of themselves, learn what teachers expect, practice hitting valued targets and decide whether to care" (pp. 208-209).
However, during this research, the teachers had not progressed to the point where they were secure enough that they could allow the students to share in the development of these criteria and scoring rubrics. This "next step" is an important and crucial one for the teachers to pursue. By creating the criteria and the scoring rubrics with the teachers, the students take responsibility for their learning and become empowered.

In a few other situations the students thought that the teachers did too much "hand holding," and did not give the students enough responsibility. All the team teachers are warm, caring individuals in a small school where everyone knows everyone, and they have many opportunities to guide and help. According to McDonald (1993), this "warm attachment" is an asset. But this helpfulness is not always perceived as a desirable quality by everyone. One senior student commented:

They [the team teachers] have a tendency to help you out, kind of walk you through your high school career.... But I don't feel this is preparing me as well for college. ... I feel sorry for the kids here ... when they hit college. They won't have a teacher saying, "here, this is what you missed yesterday when you were gone, and this is what you need to do to get extra credit." (SPI1, 5-95)

During the assessment course, Dr. Lucas had noted the team's hesitancy to let the students fail. The teachers may need to consider that, at this level, as the students are preparing to enter the "real-world," the students should assume more responsibility for their own success or failure. There seems to be
a fine line between coaching, helping, and preventing students from accepting responsibility.

Another positive outcome of using authentic assessment in the Senior Seminar relates to the exhibition. As seniors prepare for and present their projects, the teachers are identifying weaknesses in the students' knowledge and skills, and attention to these is being worked into the K-12 curriculum.

**Strengths of Authentic Assessment**

The teachers and students in this study saw the authentic assessments as learning experiences. The students recognized and realized strengths of which they were not even aware. According to the Senior Seminar students, the authentic assessments made school more meaningful, allowed them to shine in certain areas, and encouraged them to become responsible for their own learning. This is consistent with previous research which has shown that assessment has a direct impact on students' self image (Collins, 1990). An additional positive outcome is that the students indicated that they learned more from authentic assessments.

**Weaknesses of Authentic Assessment**

Weaknesses or constraints of authentic assessments as perceived by the teachers in this study are that they take more time to prepare, implement, and grade. Some students in the
Senior Seminar perceived them as harder and requiring more time and effort than traditional exams. This might be considered less of a limitation or weakness if teachers, students, and administrators, as well as parents were to change the way they think about assessment, that is, adopt the idea that assessment should not be something done to students after instruction, but should be part of instruction, as a continual process which improves student performance. Then, and only then, might the extra time spent on assessment be considered worth the effort.

**Understanding and Adopting a New Paradigm**

The literature and recent studies emphasize the importance of having the teachers understand "the philosophical and conceptual bases of the intended curricular goals" (Shepard, 1995, p. 42). The administrators in this study dealt with the teachers' philosophy of instruction and assessment by approaching this issue right up front in the assessment course. This, however, turned out to be an area of contention because the teachers felt they already had dealt with this, and too much additional course time was spent on it.

Although there is no question that teachers need to understand the philosophy behind the changes, extreme diplomacy is necessary to insure acceptance and avoid resistance and bruised egos. This study had similarities to the Colorado study (Shepard, 1995) where the researchers assumed that all the
teachers were true volunteers open to new ideas and changes, but that was not exactly the case. Some of the Colorado teachers had been "volunteered" by the principal or colleagues and were excited about learning new practices, while others were content with what they had been using. Although no one volunteered the teachers at New Albany, in some cases teachers felt the need to show the new administrators that they were enthusiastic and willing to take advantage of the assessment course that was offered to them, but some teachers were quite content with what they had been doing.

The philosophical shift has not been easy, because it has meant trying to change the way people think about school, how things are taught, how things are assessed; it has meant totally reworking how they "do school." At least "on paper," the administrators have dealt with many of these issues in the Strategic Plan (1994). According to McDonald (1993), engaging in authentic assessment and "assessment by exhibition is a systemic thing" (p. 485). Most researchers dealing with staff development and/or assessment reform agree that changes in a school must be accompanied by a compatible paradigm shift and include changes in other aspects of the curriculum and the way of doing school (Fullan, 1991; Martin-Kniep and Kniep, 1992; Shepard, 1995; Sparks, 1994). Although those in the leadership roles at New Albany recognized the need for change and created a plan which espouses a change in paradigms and a new way of doing school, in
some ways the district is still adhering to a traditional view of school.

Some of the areas which have not changed and which affect what the teachers can do and what actions they are willing to take are the district's grading and reporting system, the state's proficiency tests and the other standardized tests required by the district, and college admission requirements. The administrators at New Albany have been supportive in terms of providing opportunities and encouragement for teachers to experiment, but it has been within the confines of the current reporting and testing system. This creates frustration and problems for the teachers who must continue to prepare students for standardized, objective tests while adding the new assessments. Some schools, when introducing alternative assessments, have called a moratorium on standardized tests for a period of time while implementing the new assessments (Hiebert and Davinroy, 1993).

The mismatch of instructional and assessment practices and the reporting system also creates problems. The teachers and administrators at New Albany agree that many techniques, including objective assessments and authentic assessments, are needed to improve student performance and assess progress adequately. But these various techniques elicit much rich and descriptive information which teachers must compress and convert to letter grades. The administrators at New Albany have begun to explore options in this area by expanding the concept of
student led conferences to the high school and sending several teachers to listen to Everett Kline from CLASS (The Center on Learning, Assessment, and School Structure) concerning "the ramifications of using authentic assessments for reporting to the home" (GI3, 4-95). They should also consider other alternatives that have been proposed to bring the reporting systems into better alignment with the new assessments (Seeley, 1994; Sperling, 1994; Wiggins, 1994). Some schools have created more descriptive evaluation reports for parents (Clarridge and Whitaker, 1994) while others have devised schoolwide rubric reporting systems which correspond with their classroom rubrics (Kenney and Perry, 1994).

Another alternative is longitudinal reporting based on a student's progress toward exit-level standards (Sperling, 1994; Wiggins, 1994). Stiggins (1994) devotes a chapter to the issue of reporting to the various audiences. Besides trying to match instructional plans with assessment plans and reporting, those proposing these new systems all agree that it is vital to have a system which is "interpreted by all members of a school community in the same way" (Seeley, 1994).

Rethinking the whole system including grades, reporting of grades, and schedules will necessitate, to an even greater degree, communication and education of the parents and the public in terms of the paradigm shift and assessment literacy. Lack of communication between teachers and administrators, between
teachers and students, and among teachers, administrators, and parents in this study was a major roadblock in the beginning of the process of introducing and implementing authentic assessment into the district. A major theme which emerged during the course of this study has to do with the different perceptions of the participants within the various roles represented.

Although the administrators had knowledge of the latest theories of learning and professional development and held a constructivist view of knowledge, initially, they made many of the decisions concerning "what" was to be changed and "how" that would be accomplished. Formal introduction of authentic assessment into the district was the result of the superintendent's strategic plan for the district. Based upon the administrators' perceptions of the New Albany teachers' knowledge of assessments and what they felt these teachers should know, specific plans were set into motion. From the teachers' perceptions of what was being done by the administration in the way of the assessment course and its content, the teachers felt that they were not given credit for what they knew about learning, instruction, and assessment; they felt that their competency was being questioned. This created a situation in which what was taught, what was heard, what was learned, and what was implemented was resisted.

The recommendations for professional development based on a constructivist model of learning could just as easily be those
for classroom teaching: The general consensus calls for professional or staff development that acknowledges and works with the school culture, meets the needs of the teachers and the school, involves the teachers in the planning process, allows them to take some of the responsibility for their own learning, utilizes their knowledge and expertise in the instructional stage, while accomplishing this over an extended period of time rather than in one shot workshops. The role of the administrators and principal should be supportive but not overbearing.

This is what eventually happened in the New Albany schools. As the teachers and administrators got to know one another as individuals rather than as "other," they each came to respect the other's knowledge and attributes. Both teachers and administrators began to work together, as much as possible given their roles, to deal with the problems and issues inherent in developing and using authentic assessments. After two years, the team teachers and administrators in this study have developed a working relationship and are following much of the advice of experts in the areas of change, staff development, and authentic assessment. Now it is up to the trailblazers to guide and support the other groups: the "pioneers," the "settlers," and the "stay at homes" (Schelechty, 1993).
Where Are They Now?

Many problems have been encountered during the journey which is far from over, but both teachers and administrators seem to be addressing the main issues in a positive, optimistic, and enthusiastic manner. They see the benefits and strengths of using authentic assessment and are committed to giving it their best shot. The district has opened its doors to mentors and critical friends like myself. The administrators and teachers are willing, in most cases, to consider criticism and suggestions, to analyze what is working and what is not, and to reevaluate when necessary. Moreover, there now seems to be more collaboration, cooperation, and negotiation between the various positions or roles within the district.

Authentic assessment needs to be accompanied by a paradigm shift in order for it to work most effectively. This district is in a transition from their previous paradigm and has moved a considerable distance. On a continuum from a traditional paradigm to a constructivist paradigm, I believe they are past the halfway point.

Where Do They Go From Here?

In addition to moving authentic assessments into all the classrooms, the district's next step should be to expand the discussion beyond the teachers and administrators to include the parents and the rest of the school community. When these groups
are educated about all the changes, and the issues discussed earlier are addressed, then perhaps the paradigm shift can be completed. Until parents understand where the school is going and why the changes are being made, they, too, may resent the changes and resist them. If this happens, we may see a repeat of what happened in Littleton (Davis and Felknor, 1994) or Bensenville (Mirel, 1994). Having spent two years getting to know the administrators and teachers at New Albany, I am optimistic about the future for New Albany. I believe they understand that a paradigm shift is necessary and they are willing to make the additional changes to the infrastructure that will enable that shift, but in order to gain needed resources and support, it will also be important to expand the concept of testing through assessment literacy for parents and the general public and to inform and educate them of the positive value and long range aspects of authentic assessment.

My purpose for this study was to describe the formal introduction of authentic assessment into this district, to describe the problems, issues, and concerns of teachers during the implementation of authentic assessments, and to describe the assessment practices of four high school teachers before, during and after their formal introduction to the ideas and practices of authentic assessment. I have described the process by which New Albany administrators introduced authentic assessment into their district and have enumerated some of the problems and issues
that were encountered during that process. I have provided descriptions of how the four senior team teachers implemented authentic assessments into the Senior Seminar, what worked, what did not, some of their concerns and problems, and their attitudes about the authentic assessments they used. I have also discussed what I observed and heard in terms of what has appeared in the literature, and have pointed out areas and issues which need further consideration by the administrators and teachers.

At a time when educational reform is a prominent issue and assessment is at the forefront of the forces driving reform, it is important to observe and take note of what happens as school districts, schools, or individual classrooms take steps to initiate such changes. From these pioneers, we can see what works in various situations so that we are more knowledgeable and more prepared for similar journeys.

**Recommendations for further research**

A case study deals with one site or issue. It takes many case studies to create a full picture of a movement. There are still many areas in authentic assessment which need further research to complete this puzzle. I would like to suggest a few questions for examination by further research.

- What role can technology play in authentic assessment?
• How can grading and reporting grades be brought into better alignment with authentic assessments?

• How do we communicate to parents and the public that mere numbers aren't necessarily the best indicators of learning?

• Can authentic assessment be used on a broader scale such as statewide or is it destined for use within the classroom?

• Do students learn to think and solve problems better with the use of authentic assessments?

• Does the use of authentic assessment enhance students' ability to self-assess and improve learning?

• What is the effect of authentic assessment on motivation and the incentive to continue life-long learning when one becomes responsible for and reflective on his or her own learning?

• What implications does this have for instructional design models?

• Do elementary and secondary teachers use authentic assessments differently?

• How do students who have been assessed with authentic assessments perform in college? and in the work world?
LIST OF REFERENCES


281


*Venture Capital School Improvement Grant Application, Second Round*. (1994). Ohio Department of Education


Educational Services and Research  
New Albany Schools  
Design and Implementation of Alternative Assessment Strategies for Educators  

The Ohio State University  

There are interesting and fundamental changes occurring in how educators think about and deliver instruction in the public schools. Some of these concepts are: student as worker and teacher as coach or facilitator of learning; cooperative learning; thematically-based instruction; interdisciplinary instruction; instructional teams; etc. None of these changes in practice is more dramatic than the emergence of alternative strategies designed to assess student learning. Educators spend between 10% and 20% of their time doing student assessment. Some of the assessment is formal and standardized, while other aspects are informal and usually unstandardized. As educators, we need to be comfortable and competent with both forms. We also need to expand our knowledge of student assessment beyond where it has been for the last fifty years.

A formal educational offering is being designed and offered to the educators in the New Albany Schools. The offering is designed to: a) present the most recent information available on alternative and traditional assessment procedures; b) assist the participants in developing and adapting assessment strategies appropriate for their own instruction; and, c) facilitate the implementation and assessment of their designed strategy in their own instructional setting. Participants can work individually or in small teams.

A two credit hour offering is prepared for Winter Quarter, followed by a one to two credit hour follow-up course (implementation and assessment of the designed strategy) Spring Quarter 1994. The course will be scheduled Wednesdays from 3:30 - 6:00. The class will be held on the New Albany School grounds.

Objectives: The course will:

a. present three specific models of student assessment;

b. present the theoretical foundations and practical applications of at least four alternative assessment strategies;

c. provide information and practical experience to help participants to evaluate and interpret the standardized tests (and student scores) used in the New Albany Schools, including the tests strengths and limitations;
d. provide the necessary information and require participants to
develop a personal philosophy of student assessment, in
written form, suitable for sharing with parents and other
relevant constituencies;

e. provide the necessary information and experience to have
participants create and implement a rubric for assessing
student performance in response to an instructional unit;

f. require participants to develop an alternative assessment
strategy for use in their own instructional setting;

g. require participants to implement the alternative assessment
strategy (during Spring Quarter); and

h. require participants to assess the implemented strategy
according to a rubric which they create.

The participants should find the material interesting and motivating. It
should also enhance their repertoire of instructional skills.

Topical Outline

Jan. 5 (Wed) Models of Student Assessment, indicators of student
learning; what do they know and what can they do? Philosophy of assessment discussion and assignment.

Jan. 19 (Wed) Presentation of Alternative Assessment Strategies.

Feb. 2 (Wed) Alternative Assessment Strategies (continued).


Mar. 16 (Wed) Presentation of Cooperative Group assignment on
Standardized Test Interpretation.

An additional class meeting will be scheduled after the course begins.
The content of the additional class will focus on issues and directions for
student assessment including grading, confidentiality, and ethics.
APPENDIX B

RUBRIC FOR FIRST YEAR'S PRESENTATION
RUBRIC FOR FIRST YEAR'S PRESENTATION

Oral Presentation

A. Presentation Time frame
   Did the student use the appropriate time frame
   for verbal and media presentation?

B. Delivery
   Did the student use appropriate public speaking
   skills?

C. Organization
   Did the student logically organize and present
   the project content?

D. Extemporaneous Defense
   Did the student answer questions and defend the
   product/project?

E. Interdisciplinary Approach
   Did the student demonstrate an understanding of
   integrated skills?
APPENDIX C
TEACHERS' PHILOSOPHY
PHILOSOPHY

We believe that New Albany High School seniors, soon to assume full adult responsibility, should be offered varied, authentic educational opportunities as well as increased responsibility. Students are encouraged to go beyond the school setting, to draw upon a much larger resource base and to both extend and apply their knowledge as participants in a greater community. Through self-directed learning opportunities, students can experience the freedom and challenge to exhibit integration and synthesis of acquired knowledge by way of pragmatic demonstration. These multiple experiences, both inside the traditional school setting and outside the classroom boundaries generate benefits for an orderly and systematic transition to the world beyond secondary education. We believe that the collaborative alliances among school, community, businesses, and other resources help seniors become articulate critical thinkers, thoughtfully focused citizens, academically competitive learners, and reflective problem solvers.
APPENDIX D

TEACHERS' EXPECTATIONS OF COURSE
TEACHERS' EXPECTATIONS OF ASSESSMENT COURSE
FROM QUESTIONNAIRE

High School Teachers

• Different ways to assess students.
• How to better evaluate students progress.
• New methods of alternative assessment.
• Some alternate ways of grading my students- new ways.
• New ideas and ways to assess.
• I hoped to find out new ways to assess students.
• Get ideas on what is being done at N.A. and get ideas for things I can do. The collegiality is always important to me.
• Whether or not I am on the "right" track with classroom activities.
• Models of assessment - other than we already have; hoped to see "real" examples of assessment.

Middle School Teacher

• New ideas and useful activities

Elementary School Teachers

• As a first grade teacher & in a school that is changing to integrated curriculum I hoped to learn how to "grade" authentically my childrens work and processes of learning.
• How to identify authentic assessment activities and use them effectively.
• I hoped to learn about different methods of assessing students.
• More about authentic assessment.
• More about how to implement activities in classroom.
• To better assess students and feel more confident in doing so.
• How to assess my students effectively and discuss my students' progress with their parents.
• I hope to learn more about authentic assessment.
APPENDIX E
SECOND YEAR PROJECT RUBRIC
<table>
<thead>
<tr>
<th>I. DEVELOPMENT OF PROJECT/PRODUCT (100 Points)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIME MANAGEMENT (40 points)</td>
</tr>
<tr>
<td>Did the student meet faculty imposed and self-imposed deadlines? (20 points)</td>
</tr>
<tr>
<td>Did the student exhibit organizational skills? (20 points)</td>
</tr>
<tr>
<td>SELF DIRECTED LEARNING (30 points)</td>
</tr>
<tr>
<td>Did the student display a good work ethic? (20 points)</td>
</tr>
<tr>
<td>Did the student value the project as a learning experience? (10 points)</td>
</tr>
<tr>
<td>RESOURCES (20 points)</td>
</tr>
<tr>
<td>Did the student use appropriate resources and/or talents?</td>
</tr>
<tr>
<td>PROCESS JOURNAL (10 points)</td>
</tr>
<tr>
<td>Did the student evidence weekly written entries summarizing the project/product?</td>
</tr>
</tbody>
</table>

| TOTAL |

<table>
<thead>
<tr>
<th>II. PROJECT/PRODUCT (70 Points)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VALUE OF THE PRODUCT (10 points)</td>
</tr>
<tr>
<td>Did the student create a product which has value to other learners?</td>
</tr>
<tr>
<td>QUALITY OF PRODUCT (50 points)</td>
</tr>
<tr>
<td>Did the student create a product that was carefully crafted?</td>
</tr>
<tr>
<td>APPROPRIATENESS OF PROJECT/PRODUCT (10 points)</td>
</tr>
<tr>
<td>Did the student select an appropriate product?</td>
</tr>
</tbody>
</table>

| TOTAL |
### III. ORAL PRESENTATION (70 Points)

<table>
<thead>
<tr>
<th>Category</th>
<th>Points</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presentation Timeframe</td>
<td>10</td>
<td>Did the student use the appropriate timeframe for the verbal and media presentation?</td>
</tr>
<tr>
<td>Delivery</td>
<td>30</td>
<td>Did the student use appropriate public speaking skills? (20 points)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Did the student use appropriate delivery methods (verbal and/or media) (10 points)</td>
</tr>
<tr>
<td>Organization</td>
<td>10</td>
<td>Did the student logically organize and present the project content?</td>
</tr>
<tr>
<td>Extemporaneous Defense</td>
<td>10</td>
<td>Did the student answer questions and defend the product/project?</td>
</tr>
<tr>
<td>Content</td>
<td>10</td>
<td>Did the student incorporate knowledge, facts, data and insights gained during research?</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### IV. INDIVIDUAL WRITTEN COMPONENT (20 Points)

<table>
<thead>
<tr>
<th>Category</th>
<th>Points</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self Assessment</td>
<td>20</td>
<td>Did the student realistically evaluate his/her effort on both the process and the product?</td>
</tr>
</tbody>
</table>

**TOTAL OF PARTS I - IV**

**COMMENTS:**