CRITERION-REFERENCED ASSESSMENT AND EVALUATION 
AS SOCIALLY SITUATED PRACTICE

By

J. Karen Reynolds
B.A., McMaster University, 1984
M.Ed., Lakehead University, 1991

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF 
THE REQUIREMENTS FOR THE DEGREE OF 
DOCTOR OF PHILOSOPHY 
in 
THE FACULTY OF GRADUATE STUDIES 
Centre for the Study of Curriculum and Instruction

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
1997

© J. Karen Reynolds, 1997
In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Centre For The Study of Curriculum and Instruction
The University of British Columbia Vancouver, Canada

Date October 14, 1992

DE-6 (2/88)
ABSTRACT

In the fall of 1994, the provincial government of British Columbia introduced education policy involving an emphasis on criterion-referenced evaluation, and the reintroduction of letter grades from grades four to seven. In this dissertation I describe and interpret how one school principal, one grade 5/6 teacher, a group of grade 5/6 students, and parents of grade 5/6 students, understood the policy change. The purpose of this case study is to try to understand varying points of view that participants bring to such a policy change.

This study was a collaborative project from 1994 to 1997. "Anne," a grade 5/6 teacher, and I jointly collected and examined documents pertaining to education policy and Anne's classroom practices associated with assessment-evaluation. Data collection techniques involved participant-observation in the form of informal and formal observations, and informally-structured interviews.

A key finding of the study is that policy is ambiguous and can be understood from multiple and divergent standpoints. Anne saw policy and classroom
practices associated with assessment-evaluation as means for promoting learning in terms of metacognitive development. Many of Anne's students and their parents saw assessment-evaluation as means for generating marks and grades for exchange on competitive markets.

James Wertsch's (1991; 1995) and Pierre Bourdieu's (1990) philosophical and methodological viewpoints proved to be useful frames for interpreting findings. Within this context, I understood participants' interpretations of policy and classroom practices associated with assessment-evaluation as socially situated in settings characterized by unequal relations of power among policy-makers, school administrators, teachers, students, and parents.

My overall conclusion is that a useful approach to assessment-evaluation would provide teachers with adequate time and support to explore, discuss, and reflect upon their interpretations of assessment-evaluation policy and classroom practices within and across contexts of their work. More equitable relations of power among teachers, students, parents, administration, policy-makers, and researchers would facilitate greater communication and collaboration.
within and across interest groups. Increased communication and collaboration among people promise shared, albeit divergent, understanding regarding what counts as useful policy, and fair practices.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Abstract</strong></td>
<td>ii</td>
</tr>
<tr>
<td><strong>Table of Contents</strong></td>
<td>v</td>
</tr>
<tr>
<td><strong>Acknowledgements</strong></td>
<td>x</td>
</tr>
<tr>
<td><strong>Dedication</strong></td>
<td>xii</td>
</tr>
<tr>
<td><strong>CHAPTER</strong></td>
<td></td>
</tr>
<tr>
<td>I ASSESSMENT AND EVALUATION AS SOCIALLY SITUATED PRACTICE</td>
<td>1</td>
</tr>
<tr>
<td>The Problem and Its Contexts</td>
<td>1</td>
</tr>
<tr>
<td>Meanings and uses of key terms</td>
<td>2</td>
</tr>
<tr>
<td>My standpoint</td>
<td>7</td>
</tr>
<tr>
<td>Organization Of Dissertation</td>
<td>16</td>
</tr>
<tr>
<td>II REVIEW OF LITERATURE</td>
<td>20</td>
</tr>
<tr>
<td>Contemporary Issues in Assessment-Evaluation</td>
<td>21</td>
</tr>
<tr>
<td>Education reform</td>
<td>21</td>
</tr>
<tr>
<td>Authentic/alternative assessment</td>
<td>22</td>
</tr>
<tr>
<td>Policy Research and Assessment-Evaluation</td>
<td>32</td>
</tr>
<tr>
<td>Social and Sociocultural Research</td>
<td>40</td>
</tr>
<tr>
<td>Related Perspectives - Sociocultural Research</td>
<td>51</td>
</tr>
<tr>
<td>Related Perspectives - Social Theory</td>
<td>55</td>
</tr>
<tr>
<td>Conclusion</td>
<td>57</td>
</tr>
</tbody>
</table>
III RESEARCH DESIGN ................................................. 59
   Gaining Access ............................................. 61
   The School ................................................... 71
   Role .......................................................... 78
Methods of Data-Collection and Ongoing Analysis 89
   Documents .................................................... 89
   Informal and formal observations ......................... 91
   Informally structured interviews .......................... 92
   Field-Notes .................................................. 105
Data-Analysis .................................................. 102
   Transferability, Reliability, and Validity ............... 107
IV EDUCATION REFORM IN BRITISH COLUMBIA .......... 112
   The "Year 2000" .............................................. 114
   The New Democrats take on assessment and evaluation ............................................. 120
   The Standpoint of a Principal ............................ 133
   Conclusion .................................................... 141
V THE EXPERIENCE OF A TEACHER ......................... 146
   Anne's Background And Philosophy Of Education .... 146
   Making Sense Of The Policy Change ....................... 149
   Assessment-Evaluation Procedures and Practices ... 159
      Marking ..................................................... 159
      Criterion-referenced evaluation ....................... 168
ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to my advisor, Dr. Jane Gaskell, for her guidance, support, and encouragement throughout my doctoral studies. Jane has been my most powerful mentor. Her unique and special ways of 'being in this world' have forever touched my mind and soul.

Thanks should also be extended to Dr. Jacquelyn Baker-Sennett and Dr. Gaalen Erickson for their patience and many valuable insights as committee members. In addition, I gratefully acknowledge the "Social Sciences and Humanities Research Council of Canada," and the "Faculty of Graduate Studies" at the University of British Columbia, for their financial support.

This study would not be possible, however, without the trust, honesty, and generosity of the real "Anne Wade." Anne's commitment towards the well-being and education of her students is a wonderful tribute to her profession. To the principal and staff of "Westheights" school, Anne's students, and parents of Anne's students, I extend a sincere thank you for their support and participation. To all others who
participated at "Westheights School District" I am also indebted.

A special thanks goes to several friends who offered smiles and encouragement throughout the ups and downs of my graduate studies. To Kerrie, Irene, Bruce, Lara, Joanne, and Boo, I am especially grateful.

I must extend my deepest thank you to my mother. Without her undying emotional support, financial patronage, and faith in my capabilities, I could not have realized this dream. You're a real star Mum!
Dedication

This dissertation is dedicated to my Mum,

M. Noreen Reynolds
CHAPTER 1

Reflection on how social phenomena get defined as problems in need of explanation in the first place quickly reveals that there is no such thing as a problem without a person (or group of them) who have this problem: a problem is always a problem for someone or other. (Harding, 1987, p.6)

CRITERION-REFERENCED ASSESSMENT AND EVALUATION

AS SOCIALLY SITUATED PRACTICE

The Problem and Its Contexts

In the fall of 1994, the provincial government of British Columbia introduced education policy involving an emphasis on criterion-referenced evaluation, and the reintroduction of letter grades from grades four to seven. This thesis describes and interprets how people directly associated with one grade 5/6 class in the Lower Mainland interpreted and responded to the policy change. The purpose of this study is to try to understand the varying points of view that participants bring to such a policy change.

In this case study, the particular contexts for studying policy and classroom practices associated with assessment and evaluation are the B.C. Ministry of Education, one School Board, one school principal, one
grade 5/6 teacher, a group of grade 5/6 students, and their parents. "Anne," a grade 5/6 home-room teacher, was at the centre of pressures generated by the intersection of these contexts. She bore enormous responsibility for interpreting and responding to policy regarding assessment and evaluation, because she took the change to criterion-referencing seriously, and she, like every other teacher in the province, was expected to construct official records of student achievement in the form of report cards.

By working collaboratively with Anne, I came to understand that making sense of educational policy in classroom practice is a messy affair. Her classroom practice involved people working together, and resisting each other, in real time. Unequal relations of power among Anne, students, parents, and administration were regularly expressed as these people struggled to make sense of their experiences, and to satisfy immediate needs.

Meanings and uses of key terms. In this thesis, people use key terms differently. My use of terminology was influenced by my training and experiences as a classroom teacher, Special Education
Resource teacher, Special Education Consultant, and as a university instructor for Special Education and Pre-Service Teacher Education courses. My greatest difficulty in researching and writing this thesis has been the use of the terms, assessment and evaluation.

Although Ministry of Education policy documents distinguish between the terms, assessment and evaluation, policy discussions clearly privilege use of the term, evaluation. Ministry documents describe assessment and evaluation in the following manner.

Assessment is the systematic process of gathering information about student learning, what they know, are able to do, and are trying to do. Evaluation (is) the process of making judgments and decisions based on the interpretation of evidence of student learning gathered throughout assessment. Evaluation might be done by the teacher or the student independently or in collaboration. (British Columbia Ministry of Education, February 1994a, February, XX0246, p.21)

Although the "Kindergarten to Grade 12 Education Plan" draws from the above definition of assessment in its discussion of "evaluation and reporting," the term, assessment, is not used (British Columbia Ministry of Education, 1994i). The "Guidelines for Student Reporting for the Kindergarten to Grade 12 Education Plan" discuss assessment and evaluation in a section entitled, "Assessment and Evaluation" (British Columbia
Ministry of Education, September 1994e, p. 14). The first paragraph of the discussion involves a definition of assessment and a brief description of assessment methods. The remaining 1 and 1/2 pages focus on evaluation and, in particular, "criterion-referenced evaluation." The document describes criterion-referenced evaluation as evaluation "referenced to criteria based upon expected learning outcomes described in the provincial curriculum" (p. 14). Steps are provided for teachers to follow when using criterion-referenced evaluation. The Ministry did not use the phrase, criterion-referenced, in relation to assessment although assessment procedures were referenced to curriculum criteria.

The Ministry's emphasis on the term, evaluation, and its use of the phrase, criterion-referenced, in relation to evaluation and not assessment, appeared to place greater importance on evaluation. This led me to ask whether the Ministry's distinction between assessment and evaluation meant that assessment and evaluation were equal partners in the project of collecting information and making judgements about student progress.
My difficulty with comprehending the Ministry's distinction between assessment and evaluation, increased when I noticed that the steps for using criterion-referenced evaluation involved assessment methods (British Columbia Ministry of Education, September, 1994). This led me to wonder when assessment was not evaluation, and when evaluation was not assessment. For me, the distinction made by the Ministry is not a clear one for the collection of any student work involves deciding what criteria are going to be used to make evaluative judgements.

As previously stated, people in this thesis use the terms, assessment and evaluation, differently. People also use the phrase, criterion-referenced, differently in relation to assessment and evaluation. These people include participants, policy-makers, theorists, researchers, and myself. In this thesis, I use the terms, assessment and evaluation, in the following manner.

Assessment and evaluation represent ongoing and interconnected classroom procedures and practices for gathering information and making judgements about student progress based upon curriculum criteria and
expected learning outcomes. Within this context, assessment and evaluation are criterion-referenced because procedures and practices are referenced to curriculum criteria. When I am speaking from this standpoint, I use the expression, "assessment-evaluation."

The term, assessment-evaluation, visually signifies an interdependent relationship between assessment and evaluation. The expression uses the terms, assessment and evaluation, because teachers occasionally draw distinctions between classroom procedures and practices associated with gathering information about student progress, and classroom procedures and practices associated with making judgements about this information. Assessment-evaluation is a practice-sensitive term. It visually signifies a continuum of classroom procedures and practices with varying degrees of interconnectedness that teachers use to gather information, and make judgements about student progress based on curriculum criteria.

It is not useful to distinguish between the terms, assessment and evaluation, outside of the context of
classroom practice. Their meanings, uses, and utility must be understood from the standpoints of people directly affected.

My standpoint. My research problem is informed by my experiences as a child, teacher, and researcher. I first felt the negative impact of classroom assessment-evaluation practices when I was a student in grade 6. After the first set of report cards was sent home, my teacher announced that a new seating plan for the classroom would reflect students' grades. The student with the highest average sat in the first desk in the row by the door, the second highest average behind her, and so on, until student averages snaked up and down the rows towards the windows on the far side of the room. I was assigned a seat in the last row by the windows because of my C average. My face burned with shame as I hunched down in my new desk.

From grade six through high school and my first year of university I continued to struggle. Letter grades of Cs and low Bs made me feel inadequate as a student and as a person. I thought that high grades meant someone was smart, and smart people were popular. From the time I was eleven years old until my early
forties, I did not question the fairness of classroom assessment practices.

As a teacher I continued to believe in unidimensional and essentialist interpretations of intelligence, ability, and achievement. During my years of teaching from 1970 through the mid-1980s, the I.Q. theory of intelligence dominated the field of education. I.Q. tests were used by teachers and school psychologists to classify students as slow learners, developmentally handicapped, and gifted. As a special needs teacher, I thought that the I.Q. test defined and measured intelligence, ability, and achievement.

My uncritical allegiance to dominant theories of intelligence was reinforced by my assumption that as a classroom teacher I was not qualified to talk or write about the idea. Intelligence was the purview of school psychologists and psychometricians. Moreover, I assumed that one needed to possess above-average intelligence to engage in discussions of intelligence. I was convinced from a young age that I was not particularly intelligent.

I viewed ability as a sub-component of inherited intelligence. However, I saw ability as more
vulnerable to environmental influences and open to change. Achievement represented a student's ability to successfully complete daily assignments, classroom-based tests, and standardized tests. I understood my role of teacher as facilitator of students' abilities.

My classroom assessment-evaluation practices included teacher-developed tests, end-of-unit tests in teachers' manuals, and standardized tests. I thought that standardized tests represented the most reliable procedures for gathering information about students' strengths and weaknesses, and for making judgements about students' programs. I constructed letter grades for report cards by comparing information from individual test scores with class averages for various subjects.

Standardized tests allowed me to determine which students fell within the average range of achievement according to their grade placements. These tests included the "Canadian Test of Basic Skills," "Canadian Achievement Test," and "Wide Range Achievement Test." Four assumptions guided my support of standardized tests. First, since standardized tests represented achievement levels of a large population of
students, I assumed that standardized tests were free of local bias. Second, I assumed that what counted as average ability to learn was universal and quantifiable.

Third, I thought that knowledge was a universally recognized body of legitimate information; recognized, that is, by educated people. I decided that if knowledge counted as universally legitimate information, and if the ability to learn was represented by an innate quality of the individual, then standardized tests measured the ability to learn and achieve according to a universal standard.

Fourth, I believed that naturally occurring and innate inequalities in ability existed among individuals due, to some extent, to inherited differences in intelligence. Therefore, differential achievement among students on standardized and criterion-referenced tests was fair as long as all students had equal access to the tests. I understood equal and fair as same.

It seemed natural to me that unequal abilities should create differential performance in achievement which in turn created unequal access to educational
opportunities and employment. Stratification of employment in the workplace was a normal response to unequal abilities in terms of achievement. Because I believed that abilities and achievement could be enhanced with appropriate educational opportunities, I focused on remediation and enrichment.

My greatest commitment to maximizing ability and achievement was as a teacher of gifted students. Gifted students with whom I worked were identified as those who scored 130+ on an IQ test (WISC-R), and who demonstrated that they were two or more grade levels above age-related peers on standardized tests of achievement in reading and mathematics. Six years of training and experience in the field of gifted education reinforced my belief that ability was a sub-component of intelligence and that achievement was a product of the two.

My assumptions about assessment, evaluation, and achievement were severely challenged during the last two years of my classroom career when I worked as a Special Education teacher and Consultant for Pikangikum First Nations (1987-1989). Pikangikum is an isolated Ojibwa community in the sub-Arctic region of Ontario.
The first language of students was Ojibwa.

In Pikangikum school, students were obligated to compete for marks and grades. Teachers and administration told students that marks and grades could be exchanged for access to high school and employment. However, ancestries of poverty dominated children's lives. Unemployment, disease, alcoholism, substance abuse, sexual abuse, violence, neglect, and teenage suicide far outnumbered opportunities for employment in the Hudson's Bay store and school. Most kids dropped out of school by grade 10 to collect welfare with their parents. Suicide among children and youth was becoming a serious problem. Students were caught in the middle between the schools' interpretation of assessment-evaluation as means by which marks and grades were generated for exchange, and a community with nothing to exchange.

Pikangikum Band hired me to design a special education program for students from Kindergarten through Grade 8. Many of the 500 students had previously been tested with norm-referenced instruments, particularly in language arts and mathematics. Most children fell below norms for
average achievement. The non-Native Director of Education for the Band asked me to create assessment instruments to identify local norms of achievement for reading comprehension and computation.

I decided to define local norms of achievement by assessing how effectively students demonstrated their understanding of information they had been taught. The study involved students and teachers from two grade 1 classes, two grade 2 classes, one grade 3/4 class, two grade 5 classes, one grade 6 class, and one grade 7 class. I created and administered criterion-referenced tests to students in skills associated with reading comprehension and computation.

Average and above-average achievement levels of many students indicated to me that they needed more challenging and interesting programs. However, the results of my study were in conflict with many of the staff's judgements of their students' abilities and achievement levels. Teachers assumed that most students needed remedial assistance and would not benefit from enrichment programs.

In spite of my recommendation that many students would benefit from more interesting and challenging
learning experiences, the non-First Nations Director of Education told me that students would not qualify for enrichment or gifted programs until they demonstrated average or above-average achievement on standardized tests. Moreover, this Director conceded that I could reassign students to basal readers that matched the findings of my study. Consequently, I reassigned the majority of the students in the study to readers of greater difficulty. I wondered why this man even hired me to identify local norms of achievement when he was solely committed to standardized norms. Standardized tests privileged the content of Euro-Western standardized practices, their standards for mastery, and their methods of execution.

After re-examining standardized instruments, I decided that they were linguistically and conceptually biased against the students. English was the sole language of the tests and many of the tests' items referred to objects and experiences which were culturally irrelevant, and even in conflict, with students' lives. In a later study of the cultural relevance of giftedness in another Ojibwa community, I explored issues associated with the use of external

I could have concluded that the overwhelming evidence of students' below-average scores on standardized tests reflected substandard intelligence and ability. However, the preponderance of average and above-average scores that students received on criterion-referenced tests led me to wonder if their scores on standardized tests were symptomatic of an approach to assessment that unfairly discriminated against certain students.

I was concerned with the range of assessment-evaluation practices and their interpretations that were privileged for use in Pikangikum school, because I thought that standardized approaches were culturally biased against the students. Moreover, I was deeply disturbed by the non-Native Director of Education's irrevocable power to enforce a particular interpretation of achievement. His insistence that norm-referenced practices be privileged for use by teachers in their classroom assessment practices illustrated a disturbing case of institutionalized racism.

My experiences in Pikangikum shaped my belief that
the power to name what counts as assessment-evaluation and achievement represents the power to authorize a particular interpretation of assessment and achievement. My studies of achievement in Pikangikum and Sweetgrass led me to question previously held assumptions that intelligence and ability represented naturally occurring and innate qualities of individuals; and, that achievement was a product of these qualities. I continue to reexamine the utility of interpretations of achievement to educational practice in general, and to the practice of assessment-evaluation in particular.

I recognize that I have devoted much of this chapter to my experiences in Pikangikum. However, without these experiences I would not have seen a need for this thesis. I hope that this thesis contributes in some way to a more equitable approach to assessment-evaluation.

Organization Of Dissertation

The remainder of the dissertation is organized in the following manner. Chapter two begins with a review of literature on assessment-evaluation, including issues associated with validity and reliability in
relation to authentic or alternative approaches.

Next, I look at literature regarding contemporary issues in policy research followed by a discussion of how these issues are addressed within the context of interpretivist research traditions. The discussion moves to an examination of Wertsch's (1991; 1995) sociocultural theory and Bourdieu's (1990) theory of genetic structuralism including their contributions to an interpretivist approach to assessment-evaluation. The chapter ends with a review of literature on cognition and achievement within the context of social and sociocultural research traditions.

Chapter three begins with an introduction to "Westheights" school and Anne's grade 5/6 classroom. Following this description, I outline the nature of case study research and my data-collection and data-analysis techniques. Next, I discuss how Anne and I developed our collaborative relationship including difficulties and advantages of this type of arrangement. Considerations of validity, reliability, and transferability of the study end the chapter.

In chapter four, I look at the political contexts of the policy change from the standpoints of the
British Columbia Ministry of Education and "Grant," principal of Westheights elementary school.

In chapter five, I examine Anne's understanding of, and responses to, assessment-evaluation policy and her classroom assessment-evaluation practices. Particular attention is given to difficulties generated by Anne's interpretation of criterion-referenced evaluation and criterion-referenced letter grades. The chapter ends with my interpretation of Anne's experiences within the contexts of Wertsch's (1991; 1995) and Bourdieu's (1990) theories.

In chapter six, I look at students' interpretations of Anne's assessment-evaluation practices. I explore tensions generated by these interpretations in relation to student identity and students' relationships with teachers, peers, and parents. The chapter ends with my interpretation of students' concerns based upon Wertsch's (1991; 1995) and Bourdieu's theories.

In chapter seven, I describe and examine parents' interpretations of assessment-evaluation policy and classroom assessment-evaluation practices. I explore means by which parents understood assessment-
evaluation policy and practices including tensions generated by these interpretations. The chapter ends with my interpretation of parents' interpretations from the standpoints of Pierre Bourdieu (1990) and James Wertsch (1991; 1995).

In chapter eight, I look at the significance of findings from Anne's standpoint and within the contexts of Bourdieu's and Wertsch's theories. The chapter ends with a discussion of implications for theory, policy, practice, and research.
CHAPTER 2

REVIEW OF LITERATURE

This review of literature is meant to provide a rationale for studying assessment—evaluation as socially situated practice. The chapter begins with a review of literature on assessment—evaluation involving issues associated with validity and reliability in relation to authentic or alternative approaches.

Next, I look at literature on policy research in relation to contemporary issues regarding measurement research. Following this discussion, interpretivist, social, and sociocultural research traditions are examined to illustrate the value of case study research for studies of cognition, achievement, and assessment—evaluation.

I give particular attention to Wertsch's (1991; 1995) and Bourdieu's (1990) discussions regarding ways in which human activity is mediated in, and by, social, cultural, and institutional settings characterized by unequal relations of power.
Contemporary Issues In Assessment—Evaluation

Education reform. In a survey of teachers' classroom assessment practices throughout B.C., Bachor and Anderson (1993) found that practices frequently included classroom observations, assignments, samples of work, and tests made by teachers. They point out that teachers interpreted practices as accurate but questioned their fairness, consistency, and the time required to use them for classroom assessment.

Bachor and Anderson use the term, interpretation, to address issues associated with evaluation. They describe interpretation as a phase of assessment wherein teachers make judgements about assessment information. These judgements are based on comparisons of student performance with goals, with previous performance, and in relation to other students. They also point out that teachers found the interpretation phase of assessment difficult to describe. That is, interpretation, or the act of making judgements about assessment information, was an implicit and tacit process for many teachers.

Bachor and Anderson also explain that "Year 2000" policy was unclear in terms of how teachers should make
judgements from multidimensional and multimethod assessment data to construct letter grades. They recommend that teachers collect interrelated assessment data within a narrow period of time during which students demonstrate achievement under differing conditions. This approach promises to enhance interpretive frameworks of assessment practices.

**Authentic/alternative assessment.** Baron and Boschee (1995) describe authentic or alternative assessment as instructional and assessment practices designed to "evaluate a student's ability to use diverse academic skills to complete real-life tasks" (p. xi). Within this context, use of the term, assessment, is intimately connected to use of the term, evaluate.

Baron and Boschee (1995) explain that authentic assessment practices consist of multiple validations of students' skills and knowledge on a variety of authentic or real-life tasks over a period of time, portfolios of student work that provide systematic overviews of skill development and knowledge attainment, and assessments of individual performance on prescribed tasks.
Baron and Boschee (1995) identify four advantages of authentic assessment. First, performance outcomes are achieved when skills are tested during their application. Second, performance assessments engage students in collaborative inquiry, information-sharing, and higher level thinking skills. Third, students are active participants in their learning. Fourth, authentic assessment allows educators to teach to the test without impairing its validity.

Baron and Boschee (1995) add that a disadvantage of authentic assessment is cost. In particular, authentic assessment is labour-intensive and is therefore more expensive than electronic assessment. Furthermore, computers have limited value in relation to authentic assessment practices because the scale of a performance-based assessment rarely exceeds ten points.

Baron and Boschee (1995) suggest that a second disadvantage of authentic assessment is subjectivity. For example, marking essays is a subjective activity and therefore can not be easily accomplished by assigning a scaled score. Therefore, validity, reliability, and comparability of assessment results go
Bateson (1994), argues that proponents of authentic assessment must address issues of validity and reliability; albeit in different forms from previous practice. He explains that validity can be addressed by focusing on the content quality of performance tasks. Content quality is achieved with up-to-date evidence of how students interpret content, and through judgements of quality by subject specialists. Face validity of performance tasks is high because they are the domain of interest.

In a discussion of reliability, Bateson (1994) points out that developing quality performance tasks is time-consuming and involves specialized knowledge. Furthermore, once students have engaged in a particular task, it cannot be used as an assessment tool. He concludes that traditional concerns for reliability are outweighed by approaches that secure information regarding achievement that is sensitive to the context of learning.

Bateson (1994) suggests that the terms, authentic and performance, are contested. Within this context, there is no consensus among educators concerning
meanings and uses of the terms. He adds that many people interpret the term, authentic, as any task or technique that is not mainstream or traditional.

Defining the assessment this way neatly protects it from any questioning about important issues such as are contained in the traditional psychometric views of reliability and validity. Conversely, any method that is not deemed authentic is, by logic, not authentic, not reliable, not trustworthy, and/or not genuine. One cannot imagine a classroom teacher telling a parent, 'Well, we want to do some nonauthentic assessment to alert us to possible problems in reading.' (Bateson, 1994, p. 235)

Bateson's (1994) concern that psychometric views associated with validity and reliability must be addressed in relation to alternative assessment practices warrants a brief review of these ideas. Anastasi (1988) explains that technical validity of an assessment instrument pertains to what it measures and how effectively it does so. She describes content-related validation as involving an analysis of test content to see if it comprises a representative segment of the realm of behaviour to be appraised; achievement tests are typically evaluated in this manner.

Anastasi (1988) points out that it is difficult to construct a representative sample of the domain of behaviour to be tested. For example, test items must
not only cover chief features of the domain, but they must also cover them in appropriate proportions. In addition, test content should involve key objectives, application of principles and the interpretation of information, and factual knowledge. Content validity primarily rests upon the relevance of the examinee's test results to the domain under scrutiny.

Anastasi (1988) describes criterion-related validity as the efficacy of a test to predict performance in particular activities; aptitude tests are constructed with concern for this type of validity. Performance is examined in relation to a criterion, or measure of that which the test is constructed to predict. Within this context, predictive validity involves predictions of future behaviour. Concurrent validity involves acquiring data to diagnose current performance.

Anastasi (1988) goes on to describe construct-related validity as the degree to which a test adequately measures a construct like aptitude. An instrument is considered to exhibit construct validity when it adequately accounts for the variants in performance that it draws.
Anastasi (1988) describes test reliability as the consistency of test scores received by examinees when retested with the same instrument on different occasions. Typically, norm-referenced standardized tests include alternate forms for reexamination purposes.

Performance-based assessment has gained widespread interest among educators and members of the measurement community. Linn (1994) comments that the rationale for performance-based assessment is rooted in criterion-referenced measurement traditions. He explains that, although changes to Glaser's concept of criterion-referenced measurement have occurred, his central thesis guides thinking among members of the alternative assessment movement.

Glaser's emphasis on describing what an individual can do and on interpreting test results in terms of an individual's location on a continuum is quite consistent with the performance assessment movement. The emphasis on authenticity in performance assessment comes from a desire to make tasks as nearly like real-world criterion tasks as possible. The goal is to have assessment tasks that are valued in their own right so that gradations in performance will be interpretable in terms of what students know and can do. (p. 12)

Debate continues regarding the relationship between criterion-referenced and norm-referenced
approaches to assessment—evaluation. Linn (1994) contends that all tests can be normed and norm-referenced interpretations of performance can be assigned to criterion-referenced interpretations of performance. He adds that it is difficult to generate criterion-referenced interpretations from norm-referenced interpretations.

Criterion-referenced assessment procedures have widespread interest among educators and academics. Simpson (1990) concludes from research regarding criterion-referenced assessment in Scotland, Wales, and England that criterion-referenced assessment practices fail to improve teaching and learning. She explains that essentialist assumptions regarding ability, learning, and performance are present in criterion-referenced procedures. For example, both norm-referenced and criterion-referenced approaches sort students. Simpson is particularly concerned with consequences of assessment approaches that interpret learning as a product of individual and naturally occurring abilities.

It is considered that criterion-referenced assessment in Standard Grade failed to deliver the intended improvement in teaching and learning outcomes because it rested on exactly the same
assumptions as those which formed the norm-referenced practices of the past - that differences in pupil attainment are the natural expression of differences in pupil characteristics, e.g. their 'intelligence'....An educationally more useful form of assessment must be based on different concepts and a more useful theory of learning. (1990, p. 172)

Simpson's (1990) discussion regarding criterion-referenced assessment practices points to the interconnectedness of procedures and practices associated with gathering information and making judgements about student progress.

Dynamic assessment traditions focus on what students are capable of learning. Anastasi (1987) explains that dynamic assessment represents a variety of approaches for consolidating assessment and development of abilities.

Lidz (1987) points out that activity and modifiability are key concepts in a definition of dynamic assessment.

The examiner and learner are both active; the examiner is an active intervener who monitors and modifies the interaction with the learner in order to induce successful learning. The learner is prodded, directed, and reinforced into a role of active seeker and organizer of information. The product of the assessment is modifiability or change in the cognitive functioning of the learner, presumably positive change. Dynamic assessment, then, is an interaction between examiner-as-intervener and, a learner-as-active
participant which seeks to estimate the degree of modifiability of the learner and the means by which positive changes in cognitive functioning can be induced and maintained. (pp. 3-4)

Within this context, gathering information about what a student can do with assistance and making judgements about this information are interconnected. An important premise of dynamic assessment is that assessment-evaluation must be continuously updated with changes in learning.

Minick (1987) explains that Vygotsky's "zone of proximal development" laid the foundation for contemporary approaches to assessment wherein examinations of what students know and can do are consolidated with examinations of what students are able to learn. Vygotsky described his approach to examining development in the zone in the following manner.

It is usual to think that indicative of the degree of development of the child's intellect is the independent, unassisted solving of the task by the child. If we would ask him leading questions or demonstrated to him how to solve the task and the child solved the task after the demonstration, or if the teacher started to solve the task and the child finished it in cooperation with other children, in short, if the child diverged however so much from the independent solving of the task, then such a solution would already not be indicative of the development of his intellect. (Vygotsky cited in Van Der Veer and Valsiner,
Vygotsky (1978; 1986) conceptualized the zone of proximal development as the interval between what an individual can do independently in a problem-solving situation, and what an individual can do in a problem-solving situation with assistance from an adult or by collaborating with more adept peers. He suggested that knowing what a child understood and could do was a good place to begin assessment. Vygotsky was disturbed by approaches that began and ended with descriptions of what children already knew and could do.

If the child is brought for consultation with the complaint that...he thinks, understands, and remembers poorly, and the psychologist offers a diagnosis of "a low level of mental development,"...nothing is explained, nothing predicted, and the psychologist is no more able to provide any practical assistance than is the physician who offer the diagnosis of a cough. (Vygotsky cited in Lidz, 1987, p. 135)

Within the context of Vygotsky's ideas, a primary goal is to determine what a child is capable of learning with assistance.

Contemporary issues in assessment-evaluation focus on the relationship between education reform and classroom assessment-evaluation practices, the role of validity and reliability in relation to
authentic/alternative approaches, the authenticity of performance-based assessment tasks, the relationship between criterion-referenced assessment and norm-referenced assessment, and the relationship between criterion-referenced assessment and teaching and learning. People create these issues. Who gets to create them and why represents the focus of the next discussion.

Policy Research and Assessment-Evaluation

Policy-makers have the power to name, interpret, and mandate key concepts and related terminology in the field of assessment-evaluation. The power of policy-makers to decide and mandate what counts as knowledge via assessment-evaluation can be understood as "symbolic power" (Bourdieu 1986a; 1989d). Bourdieu describes symbolic power as the power to produce and legitimate interpretations of social reality.

The power to conserve or to transform current classifications in matters of gender, nation, region, and, and social status, and this through the words used to designate or to describe individuals, groups or institutions. (p. 23)

He adds that possession of "symbolic capital" is a prerequisite of symbolic power. More specifically, symbolic capital represents a form of authority wherein
people who have achieved suitable recognition are in a position to impose perceived understanding. Bourdieu (1989d) explains that because individuals and groups continuously struggle for the power to introduce and impose meaning, no one has a monopoly over symbolic power.

Bourdieu's theory of power or capital and Harding's (1991; 1993) theory of "standpoint epistemology" have helped me to conceptualize my research as political engagement. In her discussion of standpoint epistemology, Harding explains that research plays an important role in the politics of knowledge production.

Standpoint epistemology sets the relationship between knowledge and politics at the center of its account in the sense that it tries to provide causal accounts - to explain - the effects that different kinds of politics have on the production of knowledge. (pp. 55-56)

Harding (1993) cautions researchers that some socially situated locations such as people from socio-economically oppressed groups are more objective starting points for inquiry than others. Furthermore, researchers who examine policy from a privileged position, without interrogating the position itself, severely limit their access to critical understanding
Knowledge claims are always socially situated, and the failure by dominant groups... to interrogate their advantaged social situation and the effect of such advantages on their beliefs leaves their social situation a scientifically and epistemologically disadvantaged one for generating knowledge....So one's social situation enables and sets limits on what one can know; some social situations - critically unexamined dominant ones - are more limiting than others in this respect, and what makes these situations more limiting is their inability to generate the most critical questions about received belief. (Harding, 1993, pp. 54-55).

Like Harding (1993), Gaskell (1988) sees a need for policy research that addresses concerns of groups that are under-represented or absent in political debate. Gaskell points out that issues are clarified and new questions generated for consideration in the political process when research begins from the standpoints of interest groups. She contends that democratic policy research explores issues from multiple and even conflicting standpoints.

Researchers can explore issues from many potentially conflicting points of view including that of the student, the parent, the teachers' federation, or the government agency, to mention only a few. It can clarify, legitimate, and expand the political agendas these groups start with instead of simply solving problems for them....One role of a controversial study is to highlight issues of concern to children and introduce them into the debate. (pp. 413-414)
Darling-Hammond (1996) shares Harding's and Gaskell's concerns regarding democratic policy research. She proposes that problems and possible solutions are clarified when researchers and participants work collaboratively to explore issues associated with educational policy and practice.

Partnerships for what we might call dilemma-ridden research... often feature hyphenated roles for researchers, teachers, and policy-makers who are doing policy, school reform, and teaching as well as looking at it.... moving beyond a world in which those who think and plan are separated from those who teach and do the work; ... working to understand, schooling, teaching, and change by engaging in the work as well as by studying it and by creating collaboratives for democratic work and action. (Darling-Hammond, 1996, p. 15)

Moss (1996) addresses assessment in terms of educational measurement. Within this context, she contends that collaborative inquiry has been absent from traditional measurement research. Moss adds that developers have traditionally created interpretations of assessment for the people who use them; a top-down approach to assessment that excludes people from constructing and even challenging interpretations that directly affect their lives.

Those who are assessed or who use the assessment information often have no ready access to reconstruct or challenge these characterizations that may have an impact on their lives (p. 24).
Moss (1996) further explains that researchers confer status on knowledge claims with naturalist interpretations of validity, reliability, and objectivity; interpretations grounded in assumptions that direct engagement with a reality unmediated by human experience is feasible. She is particularly concerned that naturalist interpretations of validity, reliability, and objectivity obviate the role of context in peoples' understanding of, and responses to, social reality.

Like Moss (1996), Hawkesworth (1989), distrusts naturalism's claims to an unmediated reality. She extends her concern to the impact of foundationalism on studies of cognition.

Critiques of foundationalism have emphasized that the belief in a permanent, ahistorical, Archimedean point that can provide a certain ground for knowledge claims is incompatible with an understanding of cognition as human practice. (p. 15)

Harding (1993), Gaskell (1988), Darling-Hammond (1996), and Moss (1996) favour studies of how meaning is produced and received within and across contexts. Moss points out that interpretations of assessment policy and related practices have serious political consequences in terms of who speaks for whom, and whose
Delandshere and Petrosky (1994) paint a picture of evaluative judgements regarding student progress as discourse in terms of who has the power to decide what counts as knowledge. They propose that people's interpretations of assessment information are shaped by values of a particular discourse regarding what counts as knowledge.

What happens in an assessment system based on the notion that knowledge is created in language within a particular discourse and is not concerned with correct or incorrect responses but, rather, is based on the interpretation of an individual's performance in order to locate it within an overall 'valued' professional discourse? Such a conception of knowledge moves assessment toward the description and analysis of discourse, and the process or act of interpretation is central to the production and understanding of individual performance. We would add that in order for judges to agree on the interpretation of a candidate's performance, they need to understand and share the ideological stance that permeates the performance standards used in the assessment process. (pp. 15-16)

Moss (1996) suggests that interpretivist research traditions allow researchers to comprehend their own, and others', constructed meaning within and across contexts in which meaning is created and received. Because interpretivist traditions assume that meaning is intersubjective, meaning must be understood from the
standpoints of people affected. Moss explains that the course and form that interpretation takes will depend on the dynamics of a context.

Like Harding (1993) and Hawkesworth (1989), Moss (1996) argues that there are no uncontested bases for establishing knowledge claims. Interpretations mediate, and are mediated by, the politics of knowledge production in specific contexts. She points out that research traditions differ in the ways in which they address the limitations of scientific inquiry.

Moss (1996) describes a case study of Special Education placement procedures to illustrate the importance of studying meaning in context. She explains that the researcher found significant differences between a school psychologist's interpretation of a grade 4 student's performance, and interpretations by the student's teacher and mother. Whereas the school psychologist's discourse was decontextualized, detemporalized, and clinical, the teacher's and mother's discourse was contextualized and sensitive to long-term considerations.

The researcher concluded that the psychologist's discourse was shaped by the time that it took to
administer intelligence and achievement tests to the student, by observing the student in a testing situation, by the test results, and by his beliefs about testing. Consequently, his discourse conveyed an image of a student who was learning-disabled across subject areas. In contrast, the discourse of the mother and teacher was influenced by ongoing observations of the student within and across school and home contexts. They shared an image of the child as someone who had learning difficulties in some activities and not others. Although both the teacher and the mother withdrew their support for the referral, the student was placed in a special class for children with learning disabilities (Moss, 1996).

Moss (1996) is careful to point out that the above story is not intended to cast an unfavourable light on the role of a school psychologist within the context of the assessment process. Instead, she wants to emphasize her view that assessment interpretations transform, and are transformed by, the particular contexts within which they are produced and received. Like Delandshere and Petrosky (1994), Moss suggests that assessment interpretations are drawn from
assumptions regarding what counts as knowledge and involve important ethical and political implications.

My point is to illustrate how interpretive methods can illuminate some of the more subtle ways in which assessment choices interact with the local context....how the local context transforms and is transformed by the assessment practice....implicit in any assessment process is a value-laden vision of what education is and of what are appropriate roles for teachers, students, and other members of the community. Epistemological choices entail ethical and political consequences. (Moss, 1996, p. 23)

Important questions regarding power and representation in relation to assessment-evaluation were raised in this section. Moreover, issues associated with who has the power to define and legitimate what counts as knowledge in interpretations of assessment-evaluation, and whose interests are represented by interpretations of assessment-evaluation, are at the heart of debate. In the next section, I explore issues associated with power and representation in relation to assessment-evaluation within the contexts of social and sociocultural research traditions.

Social and Sociocultural Research

I constructed the interpretational framework for this study from the theoretical and methodological

Bourdieu's (1990) theory of "genetic structuralism" and Wertsch's (1991; 1995) interpretation of "sociocultural theory" share an assumption that human action is dialectically constructed and cannot be decontextualized. Bourdieu (1990) suggests that interconnections among peoples' actions and objective structures are understood by focusing on how people use cultural artifacts and conventions as means to compete for capital in competitive fields. Like Bourdieu (1990), Wertsch (1991; 1995) suggests that interconnections among human action and context are understood by focusing on contexts wherein people engage in goal-oriented activity by means of cultural tools or mediatational means.

Wertsch (1991; 1995) explains that his approach to
sociocultural research is deeply influenced by the ideas of Vygotsky. Vygotsky's (1978; 1986) major premise was that an individual's higher or cultural mental functions were transformed from lower or natural functions by means of engagement with others in social and cultural activities. As previously discussed, the "zone of proximal development" represented an important theoretical and methodological context for Vygotsky's concept of mediation. Within this context, insights into a child's learning potential were gained by examining how effectively he or she solves problems with the assistance of more adept peers or adults. The goal of this approach was for a student, with assistance, to learn how to demonstrate understanding of adult-defined performance criteria.

"the zone of proximal development": the place at which a child's empirically rich but disorganized spontaneous concepts "meet" the systematicity and logic of adult reasoning....the depth of zo-ped varies, reflecting children's relative abilities to appropriate adult structures. (Kozulin in Vygotsky, 1986, p. xxxv)

Vygotsky (1978) saw the "zone" as a tool through which intrapersonal development could be comprehended.

Psychological tools and procedures for shared communication filled the role of mediators in

Psychological tools were understood as intentional movements or gestures, language and sign structures, memory enhancing strategies, and decision-making procedures. Vygotsky explained that mnemonics operated as a psychological tool in memory enhancement. For example, when the brain's natural ability requires assistance to connect one event with another event, an artificial psychological tool, like a written note, mediates the connection.

Like Vygotsky (1978; 1986), Wertsch (1991; 1995) is concerned with the mediated nature of human mental functioning. He describes the relationship between mental processes or human action, and cultural, historical, and institutional settings as dimensions of mediated action.

When action is given analytic priority, human beings are viewed as coming into contact with, and creating, their surroundings as well as themselves through the actions in which they engage. Thus action, rather than human beings or the environment considered in isolation, provides the entry point into the analysis.... (1991, p. 8)

Wertsch (1995) portrays mediated action as "action interpreted as involving an irreducible tension between mediational means and the individuals employing these
means" (p. 64). In addition, action can be internal and external, and can be carried out by individuals and groups. "Mediational means" can be understood as language and tools that bridge peoples' actions and settings within which they are used. Vygotsky (1978; 1986) and Wertsch (1995) see internal and external action as dialectical. Wertsch contends that studies of human action must be situated in contexts of time, space, and relationships among individuals.

A fundamental issue to be addressed in analyzing action, then, is how several moments of its organization are defined and understood as being involved in a complex dialectic.... Obviously the kind of thinking needed in this formulation must be based on a willingness to deal with multiple, simultaneously acting influences and the dialectics among them. (pp. 62-63)

Key to understanding the role of power relations in Wertsch's (1991; 1995) interpretation of sociocultural theory is his observation that mediational means facilitate and hinder human action.

We can never 'speak from nowhere,' given that we can speak (or more broadly, act) only by invoking mediational means that are available in the 'cultural tool kit' provided by the sociocultural setting in which we operate.... (p. 25)

Wertsch (1995) quotes Dewey, whom he describes as a "progenitor of sociocultural research," to illustrate how limitations associated with various forms of
mediation are usually recognized only in hindsight (Wertsch, 1995, p. 25). Dewey made the following remarks in 1938.

When we look back at earlier periods, it is evident that certain problems could not have arisen in the context of institutions, customs, occupations and interests that then existed, and that even if...they had been capable of detection and formulation, there were no means available for solving them. If we do not see that this conditioning, both negative and positive, exists at present, the failure to see it is due to an illusion of perspective. (Dewey cited in Wertsch, 1995, p. 25)

Wertsch (1991), and Wertsch and Rupert (1993), explain that understanding the situatedness of mediational means is key to understanding the relationship between means and action. Wertsch (1991) is careful to point out that mediational means do not always increase productivity. My favourite example of a mediational means that was not designed to enhance efficiency, is his story about the "QWERTY" typewriter keyboard. Wertsch begins by saying, "...let us consider the functioning of an object that is mediating my action as I write, the keyboard of a personal computer" (1991, pp.34-35). He adds that contemporary keyboards have their origins in the work of Christopher Latham Scholes.
In 1872 Scholes arrived at a layout representing a compromise between several demands. One had to do with the mechanics of typewriter keys. Early versions of his machines were 'slower' than typists' fingers, so the keys constantly jammed. Sholes' solution was to redesign the keyboard in an effort to slow the typist down. As William Hoffer (1985) has reported, the most common letters - E, T, O, A, N, I - became widely distributed, frequent combinations such as ED were arranged so that they had to be struck by the same finger, and the typist was required to use the weaker left hand 57 percent of the time. The familiar "QWERTY" keyboard that resulted was thus specifically designed to insure a kind of inefficiency. (Wertsch, 1991, p. 35)

Wertsch (1991) goes on to explain that even though a more efficient keyboard was created in 1936 by August Dvorak, the QWERTY keyboard maintained its dominance. Indeed, although older versions have been replaced by typewriters with electronic keyboards and personal computers, peoples' activities continue to be mediated by QWERTY and Dvorak keyboards. He reassures readers that a QWERTY keyboard can be changed to a Dvorak keyboard with a set of caps and a simple program.

Wertsch (1995) refers to pole-vaulting as an example of mediational means that tended to increase efficiency. He explains that a change from wooden to fibreglass poles increased vaulting performance. This example illustrates how changes to mediational means change actions of individuals participating in
activities. Wertsch (1993) adds that some people are more successful using some mediational means than others during various activities. He contends that questions regarding who decides what mediational means are used during activity are political in that they are configured by issues of power, value, and authority.

Wertsch (1995) acknowledges that sociocultural interpretations of mediated action have, in part, been influenced by Bourdieu's concept of "habitus." In his theory of "genetic structuralism," Bourdieu (1977) describes habitus as a mediating construct that links agency and structure. He explains that individual and group habitus involve dispositions that guide the production, reproduction, and structuration of practices within and across fields. Fields represent arenas of unequal power relations among people competing for capital or the major stakes in a field.

Bourdieu (1986a) suggests that various forms of capital mediate human activity. He classifies the forms as "cultural," "social," "economic," and "symbolic" capital. He describes "cultural capital" as legitimate knowledge manifest in dispositions, cultural goods, and qualifications, "social capital" as group
membership and valued relationships, "economic capital" as financial resources, and "symbolic capital" as power to generate "recognition" or legitimacy from "misrecognition" or illegitimacy (Bourdieu, 1990). This process can be understood as a form of hegemony wherein individuals willingly participate in practices that negatively discriminate against them.

Bourdieu (1974) describes "giftedness" as an ideology and set of practices in schools that transform cultural and social capital into institutionalized cultural capital. By awarding allegedly impartial qualifications ... for socially conditioned aptitudes which it treats as unequal 'gifts', (the school) transforms de facto inequalities into de jure ones and economic and social differences into distinctions of quality, and legitimates the transmission of cultural heritage.... The ideology of giftedness, the cornerstone of the whole educational and social system, helps to enclose the underprivileged classes in roles which society has given them by making them see as natural inability things which are only a result of inferior social status, and by persuading them that they owe their social fate (which is increasingly tied to their educational fate) to their individual nature and their lack of gifts. (pp. 115-116)

Bourdieu (1992) sees forces involved in determining the major stake or capital in a field as structuring procedures and actions by which people produce, obtain, and convert capital. Similar to
Wertsch's and Rupert's (1993) interest in forces that shape mediational means, Bourdieu (1992) is interested in forces that shape means by which capital is produced, distributed, and exchanged in a field.

The forces that are active in the field — and thus selected by the analyst as pertinent because they produce the most relevant differences — are those which define the specific capital... (capital) confers a power over the field, over the materialized or embodied instruments of production or reproduction whose distribution constitutes the very structure of the field, and over the regularities and the rules which define the ordinary functioning of the field, and thereby over the profits engendered in it.... The strategies of agents depend on their position in the field, that is, in the distribution of the specific capital, and on the perception that they have of the field depending on the point of view they take on the field as a view taken from a point in the field. (Bourdieu, 1992, p. 101)

I detect a similarity of purpose between Wertsch's (1991; 1995) interpretation of mediational means, and Bourdieu's (1986a) interpretation of capital. That is, mediational means and capital are bridging concepts in their interpretations of the relationship between agency and structure as dialectical. From Wertsch's standpoint, mediational means bridge peoples' actions and cultural, social, and institutional settings. From Bourdieu's standpoint, forces involved in defining the major stake or capital in a field structure means and...
practices by which people generate, acquire, and convert capital.

Bourdieu (1990) uses a game metaphor to illustrate relationships among habitus, practice, field, and capital. He describes habitus as a "feel for the game" that enables people to act in correspondence with the rules of a particular field, while enabling some players to use strategies that bring higher scores than others. He points out that limitations of players' actions reflect limitations of the game. Bourdieu's interest in limitations placed by social structures on agency, resembles Wertsch's (1991; 1995) interest in forces that shape mediational means and ways in which mediational means limit action.

Bourdieu's (1990) view that action must be understood in terms of a dialectical relationship between agency and social structure, is similar to Wertsch's idea that action must be understood within the context of people using mediational means during purposeful activity within and across cultural, historical, and institutional contexts. Although they use very different language to communicate their ideas, Bourdieu (1990) and Wertsch (1991; 1995) share the same
political project. Their theories offer frameworks for examining how dialectical interpretations of agency and structure change, change with, and are changed by the politics of knowledge production in fields of education, politics, and the economy.

Bourdieu's (1990) and Wertsch's (1991; 1995) assumptions that knowledge and learning are socially situated within settings characterized by unequal relations of power among people, helped me to understand how an approach to assessment-evaluation as socially situated practice might contribute to more equitable assessment-evaluation policy and classroom assessment-evaluation practices. In the final sections of this chapter, I look at related perspectives regarding sociocultural and social research.

Related Perspectives - Sociocultural Research

Like Wertsch (1991; 1995), Rogoff (1989) sees learning, or cognition, as dialectically situated. She explains that cognitive development involves "development of skill with cultural tools through participation and communication with more skilled partners" (p. 8). Rogoff's sociocultural approach to cognition involves different yet mutually constituted
levels of cultural, interpersonal, and intrapersonal sociocultural activity.

Like Rogoff (1989), an emphasis on the dialectical relationship between cognition and context frames Lave and Wenger's (1991) theory regarding situated learning. From their studies of apprenticeship among Yucatec midwives, Vai and Gola tailors, naval quartermasters, meat cutters, and nondrinking alcoholics, Lave and Wenger interpret learning as situated activity within communities of knowledge and practice. They explain that learning takes place when a person's intentions to learn are engaged and the meaning of learning is configured through the process of becoming a full participant in a sociocultural practice. This social process includes, indeed it subsumes, the learning of knowledgeable skills. (p. 29)

Lave (1988) argues that human cognition is a complex sociocultural and historical phenomenon. She suggests that cognition is situated in human practices and is "...distributed - stretched over, not divided among - mind, body, activity and culturally organized settings (which include other actors)" (Lave, 1988, p. 1).

Baker-Sennet, Matusov, and Rogoff (1992) describe children's playcrafting as a sociocultural process.

I assume that cognition, knowledge, ability, intelligence, achievement, evaluation, and assessment are socially, culturally, historically, and institutionally situated. Within this context, they are culturally-valued ideas and practices that are socially situated in unequal relations of power among people engaged in purposeful activity.

For example, intelligence represents socially and culturally-valued high-status knowledge (Sternberg, 1985; Gardner, 1986; Reynolds, 1991, 1995). In a 1986 survey, 29% of "experts" included "that which is valued by culture" in definitions of intelligence (Sternberg & Berg, 1986, p. 158). None of the respondents in a 1921 survey associated culture with intelligence.
In his theory of "multiple intelligences," Howard Gardner (1993), professor of psychology at Harvard University, explains that intelligence must be understood from the standpoints of social and cultural contexts within which it is valued.

Given the wide range of competencies...that are valued around the world, what is the nature of the mind that can give rise to a plethora of possibilities? Posing the question in this way was heterodox: it made no use of standardized tests; it focused on meaningful roles in a society rather than on abstract competencies; and it harbored a culturally relative perspective. So long as a capacity has validity in a culture, it can count as an intelligence; but in the absence of such a cultural or "field" endorsement, a capacity would not be considered an intelligence. (p. 53-54)

From past experiences as a teacher and researcher, I learned that peoples' interpretations of, and responses to, cognition, knowledge, ability, intelligence, achievement, evaluation, and assessment signified the nature of significant knowledge at a particular time, and within a specific context. I also learned that the power to name or interpret what counts as cognition, knowledge, ability, intelligence, achievement, evaluation, and assessment was not equally distributed among people. That is, some peoples' interpretations of these ideas and their applications
were privileged by assessment policy and classroom assessment practices, and not others.

Related Perspectives – Social Theory

When assessment-evaluation is examined through a lens of economic theory, learning, ability, intelligence, and achievement are understood according to their exchange value. Within this context, students might be seen as human capital. Bourdieu (1974) contends that the ideology of testing promotes norm-referenced tests as "objective" (i.e., neutral or unbiased) measures of students' abilities and achievements. He adds that classification of a student according to the conversion value of her/his cultural capital is obscured by the norming process. Conversion of capital occurs when the cultural capital of the student is tested in relation to the dominant cultural capital of the norming group. Bourdieu (1974) concludes that normative approaches to assessment privilege the interests of students represented by tests' norms.

Closely related to Bourdieu's (1974) concern that schooling's organizational processes and practices reproduce inequalities among people, are studies by
Cicourel and Kitsuse (1963), Rist (1970), and Anyon (1981). Cicourel and Kitsuse's (1963) critique of the relationship between ability and the provision of equal educational opportunity raises important issues regarding studies of assessment-evaluation. Their case study of Lakeshore High School challenges the assumption that "...equal educational opportunities for students of equal capability are insured by the use of testing devices" (p. 143). They conclude that tests commit schools to scores that are produced by these practices. In addition, social-class standing was an important factor in processing students.

Rist's (1970) single case study of one elementary school revealed that grouping practices privileged students from middle- and upper-class families. When teachers based their expectations of student achievement according to social class, they privileged the interests of students from high socioeconomic groups.

Anyon's (1981) multiple case study of the fifth and second grades in five elementary schools in New Jersey found that educational knowledge was differentially distributed according to the social
class standings of the students. Although there were similarities in their curricula, there were notable differences in the structure and the content of the knowledge which was offered by each of the schools.

Assumptions regarding what counted as knowledge, ability, achievement, and assessment-evaluation were implicit within the educational contexts studied by Cicourel and Kitsuse (1963), Rist (1970), and Anyon (1981). Overall, performance criteria and achievement standards were drawn from tests, teachers' grouping procedures, and teachers' expectations of student achievement based upon social-class standing.

Conclusion

I have crafted this review of the literature to persuade readers that interpretations of assessment-evaluation are constructed by, and construct, unequal relations of power among people. The power to name and interpret what counts as assessment-evaluation represents a key stake in the politics of knowledge production within the field of education.

Interpretive research methods offer an approach to the study of assessment-evaluation that honours the role of context in peoples' understanding of, and
responses to, education policy and classroom assessment-evaluation practices.

...interpretive methods can illustrate some of the more subtle ways in which assessment choices interact with the local context, affecting the validity of the intended interpretation and the consequences of assessment use — how the local context transforms and is transformed by assessment practice....we become better able to theorize and evaluate complex ways that assessment practices work within an educational system. And we become better able to debate and reach sound decisions appropriate to a given assessment purpose and context. (Moss, 1996, p. 23)

Interpretive methods generate important information about education policy and classroom assessment-evaluation practices from the standpoints of teachers, students, parents, and administration.
RESEARCH DESIGN

I feel really good that Anne has started to plan for my being in the class. And it's reassuring to know that I only have to get my name tag and not sign in. And I have a space in the library card box with the other teachers (Field-Note, 1.13.95).

This case study situates changes in Ministry policy regarding assessment and evaluation within the context of one grade 5/6 classroom in the Lower Mainland of British Columbia. Case study research represents an approach to empirical inquiry that examines a particular context in detail (Yin, 1994). Rist (1970) explains that a single case study of a group of students offers a detailed analysis of the intricacies of classroom life over a period of time. He adds that, "the complexities of the interactional processes which evolve over time within classrooms cannot be discerned with a single two- or three-hour observational period" (p. 75).

My choice of a case study approach for this study responds to Marshall and Rossman's (1989) view that "...the strength of qualitative studies should be demonstrated for research that is exploratory or
descriptive, and that stresses the importance of context, setting, and subject's frame of reference" (p. 46). In addition to an emphasis on context and participants' frames of reference, this study is qualitative because findings were not generated by statistical means, or other modes of quantification (Strauss & Corbin, 1990). My use of participant-observation techniques characterizes the research as ethnographic (Spradley, 1980; Hammersley & Atkinson, 1983).

I decided to situate the study in a grade 5/6 classroom because the intermediate division was the focus of a change in policy regarding letter grades. Beginning in the school year 1994-95, grades 4-7 teachers were instructed to use structured written comments and criterion-referenced letter grades to report student achievement (British Columbia Ministry of Education, 1994e). Although letter grades were mandatory, school districts decided how they were communicated to parents. Previously, "Year 2000" policy (British Columbia Ministry of Education, 1989b) encouraged teachers to write anecdotal reports regarding student progress. Letter grades were
Grade 5/6 students were experienced with the process of schooling and were conversant in the language of assessment. They were equipped to discuss their standpoints concerning classroom assessment-evaluation practices during informally-structured interviews.

**Gaining Access**

The study involved several stages of access negotiations. It was first necessary to locate an intermediate teacher and principal who were willing to participate in the study. Then, I needed to obtain the participation of students and their parents. Finally, I had to negotiate the participation of other teachers.

I began negotiations in the fall of 1994 by presenting an overview of my study to several faculty members at the university. I asked if the professors knew of a teacher or school that might be interested in my research. After several unsuccessful attempts, I met with the Director of the Teacher Education Centre who told me that he was involved with a group of intermediate teachers in the Lower Mainland who were interested in assessment-evaluation. We agreed that he
would mention my study at their next Teacher-As-Researcher meeting.

Several days later the Director notified me that the "Teacher-As-Researcher" group had invited me to participate in an upcoming meeting. The meeting focused on the implementation of policy regarding criterion-referenced evaluation within the context of the science curriculum.

During the coffee-break I met Anne, a grade 5/6 teacher. We shared our reflections on the workshop and agreed to meet informally to discuss my study. When the meeting was adjourned, the Intermediate Consultant invited me to attend an all-day professional development workshop on November 3, 1994 at the District Office. The meeting focused on implementation of policy regarding criterion-referenced evaluation in relation to classroom practices.

I visited Anne during the last week of October to discuss my research. We met informally in her classroom after the students had been dismissed. Anne said she was interested in the study because she wanted to explore issues associated with criterion-referenced evaluation in her classroom practice. During two brief
conversations the following month, Anne reiterated her interest in our working together.

During the first week of December, I received a letter from Anne's School Board stating that my study had been approved for implementation in schools within Anne's school district. Upon receipt of this information, I decided that it was time to talk to Anne and her principal, "Grant." I knew that School Board approval to conduct the study in Westheights' schools did not guarantee access to Anne's school. This meant that I needed to gain access to conduct the study from the school principal. Moreover, I understood that Anne could not participate in the research without his permission.

I was aware from past experience that a principal had the power to make the final decision concerning who gained access to a school. I also assumed that a teacher could choose to close access to her classroom through the authority of her principal. I decided that access negotiations to Anne's school would be facilitated by meeting with Anne and Grant to discuss in detail the nature and conduct of my study.

Our meeting was held in Anne's classroom on Dec. 6
1994 after 3:00 p.m. dismissal. Grant, Anne, the Director of the Teacher Education Centre, and myself were present. During the meeting Grant explained that Anne's participation on the "Teacher-as-Researcher" committee enhanced her understanding of the policy change regarding classroom assessment-evaluation practices. He added that the policy change was new for everyone and he did not have the answer regarding how it should be interpreted and implemented by teachers. Grant suggested that the major policy change for teachers was the shift from norm-referenced letter grades to criterion-referenced letter grades. He added that a great deal of work and time was required of many people to interpret and implement the change.

Grant said that he could not guarantee the participation of the Physical Education teacher and Music teacher in the study because they were very busy, and the Music teacher was new to the school. He suggested that I explore the level of interest of teachers at some point during the data-collection period, and that I spend as much time in the school as possible so that teachers could get to know me. Grant invited me to attend a staff meeting regarding
criterion-referenced evaluation on January 31, 1995; one of the school's professional development days.

We all agreed that I would begin visiting Anne's classroom in January 1995. We also agreed that I would consult with Anne and Grant to determine how and when to approach students, parents, and teachers concerning their participation in the study. Anne said that she would act as liaison between myself and the principal.

I asked how it was possible for me to be in Anne's classroom without consent from students and parents to participate in the study. Grant replied that I would be a resource person in the grade 5/6 classroom. He said that he was confident I would be useful to Anne because I was an experienced elementary school teacher. Grant ended his comments by saying that he expected me to sign in at the office each day, and wear a visitor's tag at all times.

As I left the meeting, I reflected upon the complexities of access negotiations. Different people had different stakes in my gaining access to Anne's classroom. Not only did I have to continually figure out who these people were, but I also needed to understand how they were connected. I also wanted to
be sensitive to the interests of everyone, and at the same time maintain the integrity of my research. Negotiations were intense because stakes were high. Assessment-evaluation was a sensitive issue due to its link to accountability.

With the first stage of negotiations complete, I had gained the approval of the School Board, principal, and a teacher to observe in the grade 5/6 classroom. Getting consent from parents and students was the next step.

Anne and I agreed that I needed to spend time in the classroom before sending consent forms home for interviews. She explained that students needed time to get to know me. Parents needed time to learn about me through their children's perceptions. We also agreed that students would probably join the study if their friends joined, and if they liked me.

Liking me reflected a complex and ongoing process of mutual respect and trust among students, Anne, and myself. The process developed within and across shifting contexts of time, space, and relationships. Contexts consisted of ongoing interactions among Anne, myself, and students in the classroom, school library,
Special Education resource room, computer room, and on field trips.

On February 7, 1995 Anne and I sent a letter home with students inviting parents to an informational meeting regarding the study. Consent forms for interviews were attached to the invitations. The meeting was scheduled for the evening of February 21, 1995 in the grade 5/6 classroom.

I brought enough cookies, juice, tea, coffee, milk, and sugar to the meeting for thirty people. Anne supplied a large coffee urn. Two mothers attended the meeting. They asked Anne and me how we got to know each other and how I became interested in this research. They also asked if I planned to do a comparative study that would involve a school in a lower socioeconomic context. As we ended the meeting, one woman told Anne and me that we should not judge the degree of parental support for the study by the evening's poor turnout (FN, 2.21.96).

Consent forms were returned to school within several weeks of their distribution. All of the students agreed to participate. Three families of parents refused to take part. When I shared these
results with Anne she suggested that, by telephoning the parents who declined, I could address any concerns they had. I appreciated Anne's recommendation because I would not have contacted parents without her approval.

The first parent I contacted told me that she did not want to be interviewed because her English was poor. I reassured her that I did not perceive her English to be a problem. She replied that although she would agree to a telephone interview, her husband was too busy to take part.

The second parent I contacted agreed to an interview after I reviewed the purpose and methods of the study. However, she added that her husband was too busy to take part. The third parent, a single mother, told me that she did not want to be taped. When I told her that taping was not necessary, she agreed to an interview. With parent negotiations completed, I obtained verbal confirmation from each student concerning their willingness to be interviewed.

Seeking Anne's approval to talk to parents was more than an act of courtesy on my part. It also demonstrated respect for Anne as the teacher of the
When I was teaching, I expected guests in my classroom to know their place. I anticipated that Anne expected the same of me. My ability to effectively exercise tact with teachers, students, parents, and administration was contingent upon my understanding of School Board, school, and classroom protocol. When I was not sure of protocol I asked Anne for guidance. Most times I was guided in my talk and actions by past experiences as a teacher.

With stage two of negotiations finished, I reflected upon the methods I had used to solicit participants. I decided that the parent meeting with two people in attendance had not been a useful strategy. Parents told me during interviews that their decision to participate was based upon their child's perceptions of me. Clearly, my access to parents was strengthened by working collaboratively with Anne and students over a period of time.

Getting consent from other teachers directly involved with the class represented the third and final stage of access negotiations. On separate occasions I met with the Special Education Resource teacher, the two Learning Assistance teachers, the Music teacher,
and the Physical Education teacher to discuss the purpose and conduct of the study. As a result of the meetings all teachers agreed to be interviewed.

A total of fifty-three people had joined the study. This number consisted of five female teachers, one male teacher, one male principal, eleven female students, thirteen male students, twenty-three mothers, and three fathers. Although all of the families were represented by at least one parent, there was a noticeable lack of participation by fathers. When fathers did participate, both parents were involved in the study. In all three of these cases the focus of their involvement was a son.

Throughout the data-collection period, I tried to exercise tact and courtesy in my relationships with teachers, students, parents, and administration. For example, one afternoon several students asked me if they could listen to rock music while they worked. The girls explained that Anne had allowed the boys to play rock music. I told them that I would ask Anne.

When I approached Anne with the girls' request she heatedly explained that there was no place for rock music in a school. I quietly returned to my chair.
When I looked in the direction of the girls, their faces were buried in their assignments. The room was so quiet you could have heard a pin drop. I decided to retreat to the washroom.

Heading down the hall I felt a tug on my arm. One of the girls had followed me. She said that she and the other girls were sorry for having set me up. She went on to explain that Anne never allowed rock music in the classroom and the girls knew that Anne would not agree to my request. I told her that I respected Anne's decisions and that what happened was okay.

After school I apologized to Anne regarding my role in the rock music incident. She reiterated her position on rock music and then told me it was okay; I just had to be careful about what I did on behalf of students. We agreed that from that point on students' must refer their questions directly to Anne.

The School

While travelling to Westheights school for my first visit, I was struck by the beauty of the neighbourhood. Large one and two-story houses posed behind thick hedges and colourful gardens. Brass door-fittings and leaded windows gleamed in the autumn
sunshine.

Streets adjacent to the school were gorged with cars transporting students. Outside of showrooms I had not seen so many Mercedes-Benzs, Lincoln Continentals, Saabs, Jaguars, Porshes, and Volvos in one place.

Dogs paraded down the sidewalk to greet young family members. Breeds included dalmatian, german shepherd, poodle, golden retriever, scottish terrier, and shih-tesu; most dogs were purebreds. Some people had two and three dogs on a leash.

Designer labels like Calvin Klein were visible on some clothing. Many women wore fashionable ensembles that consisted of silk, cashmere, wool, tweed, leather, and denim. The few men that I saw wore suits and ties. Many kids wore Doc Marten shoes, Nike running shoes, and trendy outfits.

Boys favoured wearing reversed ball-caps, mushroom haircuts, baggy jeans, and oversized t-shirts. Girls wore baggy jeans, skirts, dresses, tights, and a variety of blouses, shirts, and t-shirts. They wore their hair in pony-tails, side braids, dreadlocks, and strands of beads. Nail polish appeared to be the focus of attention among girls. Nails were emblazoned with
multi-coloured polka-dots.

Some students dyed their hair red, orange, and platinum. I wondered if Kool-aid kids and Dennis Rodman had influenced their choice of colour. On several occasions, I had seen kids on Granville Street who used kool-aid powder to dye their hair red, orange, green, and purple. Magazines and television commercials regularly featured Dennis Rodman, an Afro-American basketball player, at times sporting platinum blonde hair.

Although I have seen hundreds of schools throughout my career, Westheights held a special charm. The school building towered two stories above a circular drive. Classrooms peered through mesh-covered windows. Even though the architecture gave the school an institutional appearance, Westheights did not convey a sense of confinement.

The main entrance was framed by heavy double doors. With some difficulty I opened the doors and stepped into a large foyer. Pictures of students and teachers decorated the lobby. A tarnished metal plaque indicated that the school was built in 1922. From one end of the hall to the other, students busily painted
on large sheets of paper. To my left I saw a sign that said, "Office." I went into the office and introduced myself to "June", the school secretary. I explained to June that I had an appointment with the grade 5/6 teacher. June directed me to sign a guest book and gave me a visitor's pass to wear. She told me to return the pass before I exited the school.

I left the office and wove my way through students who were still painting. At the end of the hall, I climbed two sets of stairs to the second floor. I saw Anne talking to a student near a classroom door. She beckoned to me to enter the classroom.

I removed my coat and sat on a chair near Anne's desk which was placed directly inside the door. She took several minutes to explain to me that there were twenty-four students in the class. Six boys and six girls were placed in grade five. In grade six there were five girls and seven boys. Anne added that several teachers were directly involved with the class. These people included the Special Education Resource teacher, two Learning Assistance teachers, the Music teacher, and the Physical Education teacher. Anne told me I might want to wander around the room.
A familiar feeling crept across me as I glanced around the room. I missed this type of environment. Most of my childhood, youth, and adult years were spent in classrooms. I was more familiar with classroom cultures than those associated with kitchens, dining-rooms, and living-rooms. The classroom was my home. I had few preconceptions about Anne's classroom prior to my visit. Past experience had taught me that despite some similarities, classroom environments differed.

I saw that Anne's classroom looked similar to classrooms wherein I had taught. A tall metal filing cabinet sat in front of her desk. Adjacent to the desk was a long table with neatly piled stacks of assignments, paintings, projects, and paper. Students approached the table to use the stapler and three-hole punch. The wall behind the table was lined with blackboards. Laminated charts that listed performance criteria for undeveloped to outstanding work were posted above the blackboards.

A large gold carpet lay on the floor in front of the long table. Blackboards panelled the wall above the carpet. A chart stand, heavy with sheets of
writing, stood on one corner of the rug. There were several chairs in front of the blackboards.

On the opposite side of the room was a wall of windows. An enclosed heater rose several metres to meet the glass. Baskets of painting and science materials lined the windowsills. Projects and art activities were arranged on two long tables in front of the windows. Rows of microscopes filled one corner of a table. Boxes of classroom materials were stacked underneath. An overhead projector sat in front of the tables.

The only official storage space in the classroom was a two-door wooden cupboard that was located in the corner of the room across from Anne's desk. The cupboard was stocked with chart paper, coloured construction paper, white drawing paper, and newsprint. Storage space was limited.

The wall furthest from the classroom door was equipped with hooks for students' coats and hats. A row of dictionaries was visible on the shelf above. Nearby, a large wooden display unit held various novels.

Students' desks were arranged in groups of four
around the room. Boys sat with boys. Girls sat with girls. Anne told me that students chose where they wanted to sit. Many sat with their friends. Students occasionally left their desks and moved quietly about the room to talk to Anne, sharpen pencils, take more writing paper, or place used paper in the recycling box.

When the dismissal bell rang at 3:00 p.m., I thanked Anne for the visit and prepared to leave. She explained that from 3:00 p.m. to 3:30 p.m. she reviewed homework books and supervised students while they completed work. After 3:30, Anne marked assignments and planned for the next day.

Grant, the school principal, told me that six hundred and seventy-five pupils were registered in twenty-six classes from kindergarten to grade seven at Westheights school. Forty students were registered in two ESL classes. Seventeen students were classified as learning disabled.

He explained there were thirty-three and one-half full-time equivalent teaching staff. These people included one librarian, one school-based resource teacher, two learning assistance teachers, three
special education assistants, one physical education teacher. The remaining personnel were classroom teachers. The intermediate division consisted of fifteen classes. One hundred and twenty-seven students were registered in grade five. One hundred and fourteen students were registered in grade six.

Role

A collaborative approach framed my role as a participant-observer in the classroom. From past experiences as a teacher, I assumed that my credibility with Anne and her students depended upon my being useful in the classroom. Although I needed to find out what counted as useful from Anne's standpoint, I had some ideas of my own to offer. For example, I could help prepare materials for lessons such as art, assist with clean-up, hang student work on walls, and gather resources from the library.

Establishing trust was another matter. I needed to understand how Anne interpreted trust. I initially assumed that trust meant keeping promises. Anne explained to me that trust grew out of honesty and dependability. We promised to tell each other about any concerns we had at an appropriate time. I agreed
to be at school when I said I would be there; and on time. Credibility and trust were mutually constructed. Without trust there was no credibility. Without credibility there was no trust.

Anne and I decided that I would begin my work with her by marking students' assignments, preparing materials for science and art classes, and supervising computer and library classes. She explained that my involvement with students would increase when I successfully carried out these duties. I began with afternoon visits to the classroom on alternating days of the week. Then, I switched my visits to the mornings. As I became more familiar with classroom routines, my visits were more frequent and involved full days.

As the months passed, I gave remedial assistance to several grade six students who were experiencing difficulty with computation. One of the students was learning disabled. Anne and I team-taught social studies and creative writing classes. In an effort to increase my credibility I proposed to Anne that we implement a model I had developed for teaching research skills with children in her classroom. I had published
a booklet that describes a framework for getting kids to create and use plans to conduct and present research. After she read my booklet (Reynolds, 1990) Anne agreed to give it a try.

Anne and I team-taught students how to use ethnographic techniques to research information about their topics. For example, students conducted informally-structured and taped interviews with peers, and individuals in the community who had special knowledge of their research topic. These people involved aquarium personnel, veterinarians, medical doctors, and a Supreme Court judge. Each student obtained verbal consent to interview her/his participant. When interviews were completed, students transcribed several minutes and then colour-coded key-word notes to create themes. They used this information, along with information from books, CD-ROM, and other sources to prepare research reports and presentations. By the time I interviewed students they were comfortable with the process and understood why I used the method to collect information for my research.

Students also conducted informal observations to collect information about their research topics. For
example, students observed pet lizards, cats, dogs, horses, and sea-life at the aquarium. After they wrote notes about everything they saw, heard, and felt, they colour-coded key word notes and grouped colours to create themes. Following these activities, they shared themes from observations during small group discussions, and then used observation data along with interview information and other sources to create research reports, models, dioramas, maps, charts, games, and diagrams for their topics. Students' experiences with observations helped them to understand why I occasionally observed and made notes during class time.

Anne pointed out to me that explicit performance criteria were lacking in the student research model. She explained that students needed to know how they were evaluated when they began their research. We created sets of performance criteria for each aspect of the research project (Appendix F). Sets of criteria addressed the two planning stages, written report, four informational products, and oral presentation. Each set of criteria on the student work sheets included spaces for student self-evaluation and teacher
evaluation.

Anne and I jointly introduced criteria to the class, and then we invited students to comment on meanings of criteria, and their relevancy to research. After much discussion and negotiation among Anne, students, and myself, criteria were clarified, redefined, replaced, and even eliminated. Some stayed the same. Anne and I collaboratively assessed-evaluated students' projects while they presented their research to the class. Anne used marks and comments to construct, in part, letter grades for students' report cards.

Although students chose different topics, an example of a final project might look like the following. When Matt (Gr. 5) presented his research on the "Loch Ness Monster" to the class, he began by describing his plans. Then he shared information and several hand-drawn pictures from his written report about "Nessie." During his talk, Matt held up a brightly painted clay model of Nessie swimming in Loch Ness. He explained that this product illustrated the monster's habitat. When he finished his talk, Matt gave the class a pop-quiz involving key ideas from his
report. He ended his presentation by conducting a lively class debate on the theme, "Nessie — Fact or Fiction?"

Anne saw the modified research model as useful because it was simple and easy to use. In addition, students learned a great deal about their topics while they learned how to research. Moreover, Anne was pleased that students wanted to do more research.

Use of the research model strengthened my role as participant-observer in the classroom. Working collaboratively with Anne to modify and teach the model gave me numerous opportunities to work closely with her and students. Furthermore, the model familiarized students with the methodology and data-collection techniques of this study. While learning about their roles as researchers, students became familiar with my role.

I smile when I recall Cassidy's (Gr.5) research on jellyfish. When Cassidy showed me her subtopics for approval, I assumed that her sub-topic, sex, meant that she wanted to know how the sex of a jellyfish was determined. Her other subtopics were food, habitat, and movement. Cassidy explained that she planned to
gather information about her sub-topics from books, encyclopaedias, CD-ROM data-bases, by conducting observations of jellyfish at the aquarium, and by interviewing aquarium staff. She added that the final project would involve a diorama depicting jellyfish in their natural habitat, anatomical diagrams of jellyfish, sketches of jellyfish from her observations, a jellyfish mobile, and a written report that included detailed descriptions of her sub-topics. I approved Cassidy's topics and reminded her that I was always available to help if she encountered difficulties.

Several weeks later, Cassidy announced to me that she had encountered a serious problem with her research. After more than a decade of hassle-free research with kids, I assumed that no child could present me with an unsolvable problem. Cassidy whispered in my ear that the other kids were making fun of her sex sub-topic. In a booming voice, I announced to the class that all research was important and that we were all going to listen to Cassidy's problem.

Cassidy turned to the class and said that her problem was with sex and jellyfish. Snickers rippled throughout the room. Anne lowered her head. Without
missing a beat, I proclaimed that sexing a jellyfish must be a difficult task. Cassidy quickly turned to me and said that her problem was not with sexing jellyfish. She wanted to know how jellyfish had sex. I said no more. Cassidy asked if she might find out how jellyfish had sex from a Marine Biologist at the university. She said that when she asked her father, a heart surgeon, he replied that he did not know but Ms. Reynolds would take care of the problem.

I felt somewhat embarrassed to describe Cassidy's request in a telephone message to a Marine Biologist at a local university. Several days later, I received a detailed message explaining that little was known about the mating habits of jellyfish. The Marine Biologist ended his message by saying that Cassidy's question was interesting, and that research in this area was important. I played the tape for the class. Everyone was excited and Cassidy was a hero. She had created an important research question; one that she intends to pursue as a Marine Biologist some day.

I have many other stories that I could tell about children creating, conducting, and living research that mattered to themselves, peers, teachers, and parents;
research that created partnerships among the students, Anne, and myself. All of the students' research projects were important, exciting, and informative. Jeremy's (Gr.5) and Matt's (Gr.5) research regarding the Loch Ness monster and Big Foot sparked heated debate among students about what should count as appropriate foci for scientific inquiry.

My research interests were continually renewed by the children's enthusiasm. I also gained insights into how Anne interpreted and resolved issues associated with criterion-referenced evaluation from our work with the research model.

Anne's interest in criterion-referenced evaluation led us to jointly examine her selection and use of tests, portfolios, projects, homework books, and daily assignments to assess-evaluate student progress. I also accompanied Anne to "Teacher-As-Researcher" and "Multiple Intelligences Interest Group" meetings. Meetings were held once a month. Anne was interested in "Teacher-As-Researcher" meetings because they focused on issues associated with assessment-evaluation. By attending meetings, she hoped to gain a better understanding of criterion-referenced evaluation
in particular. Anne was interested in the "Multiple Intelligences" meetings because they focused on multidimensional conceptions of intelligence and ability. I hoped to learn how these meetings influenced Anne's understanding of assessment-evaluation.

Initially, I visited the grade 5/6 classroom on alternate mornings and afternoons so that I could observe the entire curriculum. As time went on I spent full days in the classroom. After several months had passed my visitor's tag was replaced by one that said, "Staff."

I had planned to reduce my involvement in the classroom during the interview period. I also intended to book off time to write-up observations. However, this was easier said than done. By the time I gained acceptance with Anne and students, I was an integral component of practice in process. In other words, I was needed in the classroom.

I suppose I could have reneged on some of my participation, but at what cost to my relationships with students, parents, Anne, and the principal? I decided it was not in my interests to extricate myself
from practice in process. Consequently, by May and June I was facing fourteen-hour days.

My experiences as a teacher and administrator guided my movements among teachers, students, parents, and administration. My practical knowledge of classroom, school, and School Board protocol helped me determine how to behave and what to say in different situations.

Anne and I continued to jointly teach and refine our modified student-research model. She also helped with data-analysis and gave me feedback regarding my writing. In return, I gave Anne feedback on her M.Ed. course-work. Sometimes we visited for the sole purpose of drinking coffee and chatting about family, friends, and hobbies. We also studied silk-painting after school and on weekends after her students were introduced to the technique. We still talk about the frustrations and laughter we shared while making our pictures. Anne created a lovely wall-hanging of pink, mauve, and yellow irises for her bedroom wall. My banner depicted an iridescent dragon flying past a brooding wizard. An important ingredient of our collaborative relationship has been a willingness to
give time to each other on the one hand, and to respect each other's need for privacy on the other.

Methods of Data-Collection and Ongoing Analysis

I used ethnographic techniques to collect and analyze information regarding assessment-evaluation policy and classroom assessment-evaluation practices. Specifically, I gathered and analyzed data from policy and classroom documents, formal and informal observations, informally-structured interviews, and field-notes. My approach to analyzing data involved content analysis and triangulation.

Documents. Throughout the study I collected documents associated with assessment-evaluation policy and classroom assessment-evaluation practices. I analyzed documents for key terms and major themes. Documents offered me another standpoint from which to corroborate and expand information from interview and observation data.

Policy documents organize and convey official interpretations of key assessment-evaluation terminology and related practices for school boards, school administration, and teachers. My data-base of policy documents consisted of material associated with
the "Royal Commission On Education" (1988), the "Year 2000" (1989b), and the "Kindergarten To Grade 12 Education Plan" (1994e).

Anne and I examined classroom documents for interpretations of key assessment-evaluation terminology and related classroom practices. Documents consisted of classroom tests, work samples, homework books, work sheets, student-self assessments of report cards, intermediate report card, report card inserts, portfolio guidelines, and parent feedback concerning student-led conferences.

Teachers' understanding of assessment-evaluation policy and classroom assessment-evaluation practices are influenced by school board and school-based professional development. Anne and I collected documents during Professional Development days and "Teacher-as-Researcher" meetings. Meetings focused on policy regarding criterion-referenced evaluation. We examined interpretations of assessment-evaluation terminology in workshop documents. Documents included "Criterion-Referenced Evaluation: Getting Started - A Handbook for Educators" (Kamloops, 1994), "Learning For Success" (Langley, 1994), "Assessment is....." and
"Advantages of giving students criteria..." (Doc, 11.2.94; 11.07.94), "Knowing And Showing Where The Goal Posts Are: Criterion-Referenced Evaluation" (Doc, 1.31.95), and "A.E.R. Criterion-Referenced Evaluation: The road to success is always under construction" (Doc, 2.10.95).

Informal and formal observations. As described above, informal and formal observations were ongoing from January to June, 1995. During this period I engaged in, and wrote up, one hundred and two observations. Ninety of these observations were informal, or reconstructed. Twelve of these observations were formal, or immediate. The major context for observations was the grade 5/6 classroom.

I classified an observation as formal, or immediate, when I observed and wrote about classroom events as they occurred. In terms of the participant-observation continuum, I acted solely as an observer during formal observations.

Other contexts for observations included everyday activities in the school's staff-room, halls, library, and main office, Professional Development staff meeting, intermediate division meeting, "Teacher-As-
Researcher" meetings at the District Office, Multiple Intelligences Interest Group meetings in various schools in the District, evening meeting of parents in grade 5/6 classroom to discuss the study, and telephone conversations with Anne.

My informal and formal observations culminated in a large set of written field-notes. I classified an observation as informal, or reconstructed, when I was involved to some degree in classroom activity as a participant-observer. At these times it was difficult, and even inappropriate, to write about classroom events as they occurred. Instead, I made key-word notes at recess, lunch, and immediately after school. I returned home as soon as possible to write-up complete observations from key-word notes.

Informally-structured interviews. Yin (1994) points out that the informally-structured interview is a useful technique for gathering information about an individual's standpoint regarding particular issues. This type of interview is conversational and offers participants flexibility and latitude to explore and articulate their insights and concerns.

I conducted fifty-eight informally-structured
interviews with participants during May and June of the data-collection period. Fifty-three interviews were taped. Two parents did not want their interviews taped. On one occasion, the principal and I had an unexpected meeting and I did not have the tape recorder with me.

On another occasion Anne and I spoke on the telephone. I did not feel that it was appropriate to tape this discussion. Anne and I also met casually over coffee. I chose not to tape this event either. My relationship with Anne would have been jeopardized if every time she spoke I whipped out a tape recorder. Too much surveillance can make someone feel like a bug on a pin. Instead, I took notes during and/or after untaped interviews and casual encounters.

The total number of interviews included six with Anne, twenty-three with the grade 5/6 students, twenty-three with parents, two with the Special Education Resource Teacher, one interview with the Learning Assistance Teachers, one with the Music teacher, one with the Physical Education teacher, and three with the school principal.

I interviewed one participant at a time, with the
exception of four occasions when two people took part. One occasion involved the two Learning Assistance Teachers. The other three occasions involved three sets of parents; each set consisted of one mother and one father.

Interviews ranged from fifteen minutes to two hours in length. They were scheduled before 8:30 a.m. on a school day, during recess, at lunch, during instructional periods, directly after school, in the evenings, and on weekends. I interviewed Anne at her home, in the classroom after school, and in my home. Interviews with parents and students were located in the grade 5/6 classroom, school corridors, Special Education Resource room, and psychologist's office. Several parents were interviewed in their homes. The two telephone interviews were conducted from my home.

Interview questions served as guides to discussion of major issues associated with assessment-evaluation. I encouraged participants to explore and interpret issues from their standpoints. At regular intervals in the interview process, I asked participants to clarify their interpretations and concerns. I verified my understanding of interpretations by restating them to
participants, and asking if what I said was an accurate reflection of their comments. When they answered in the affirmative, we proceeded with our discussion. When they disagreed, we renegotiated our comments until we agreed that I understood what they were saying.

I interviewed Anne on six different occasions to explore her interpretations of assessment—evaluation policy, and classroom assessment—evaluation practices. We discussed how and why she adopted, modified, and developed her interpretations over time.

Students and parents have many different understandings of teachers' assessments—evaluations of student achievement. During interviews with parents I investigated their interpretations of classroom assessment—evaluation practices in general, and with specific reference to their child. Parents reflected on why they held certain interpretations, and explained what they expected from the assessment—evaluation process and schooling.

Interviews in parents' homes were longer than interviews at the school for several reasons. The school setting presented a visually and auditorily sterile environment compared to students' homes.
Interviews in homes were occasionally interrupted by noises of children playing, telephones ringing, people coming and going, and pets begging to go out and come in. In contrast, interviews at school were undisturbed except for an occasional bell, announcement on the public address system, and students changing classes.

Social norms that guided behaviour in school differed from social norms that guided behaviour in homes. When I received a parent for an interview at school, I greeted them by shaking their hand and saying "hi." I showed this parent into the interview area and indicated a chair. Once we were settled I briefly reviewed the purpose and conduct of the study and asked if they had any questions. Then we began the interview.

Interviews in parents' homes were complicated by social codes guiding behaviour for guests. My meeting criteria for what counted as an appropriate house guest was more important than exhibiting criteria of a competent researcher. Before I left for one of these interviews, I dressed as if I were going to a School Board meeting. I freshly curled my hair, touched up my make-up, put on my dinner ring, and replaced my
backpack with a briefcase. I put on my good socks because I knew that I would remove my shoes when entering a home.

Interviews in parents' homes included time spent on light conversation and tea. Although these moments were pleasant, I was ever conscious of what I said and did; and how I did it. When a student and/or a sibling appeared, I took time to say hello and exchange a few comments. Sometimes, twenty to thirty minutes passed before I began an interview.

It was not uncommon for parents to ask what their child had said to me during an interview. Some asked me to tell them their child's pseudonym. When I reminded parents that I could not share this information, they were surprised; not by my answer, but because they had asked. They were not used to their child having a right to anonymity; and having it enforced.

Codes of conduct concerning social graces for guests differed from one home to another in terms of formality. In one setting I sipped tea from a silver tea set. In another setting I drank instant coffee from a chipped mug.
My greatest challenge was an interview with Chinese parents. Approximately twelve percent of students in the class were of Asian descent. I was unfamiliar with Chinese social norms. Both parents greeted me at the door by bowing several times. They led me into a large living-room and invited me to share tea with them. I watched how they used the Chinese tea service and mimicked their movements.

I saw an older woman standing by the door to the hall. The father told me that she was the grandmother of their children. He added that she did not speak English. I noticed that she was smiling at me. I walked over to her, took her hand in mine, and bowed my head. In Pikangikum I had learned to lower my eyes in the presence of an elder.

When the interview was finished I thanked the parents for their time and hospitality. I greeted the grandmother with another slight bow. On my way out the door I bowed to the parents. They bowed in return.

Homes were comfortably furnished. Paintings, china, sculptures, and collectibles decorated walls and cabinets. Flooring was covered with plush carpeting, ceramic tiles, and hardwood. I saw gold coloured taps
and faucets in the bathroom of one house.

Interviews with students were held before and after school, and during lunch. On several occasions Anne allowed students to be interviewed during class-time. Interviews were located in the classroom, corridors, and stairwells. Occasional noise from other classrooms and students changing classes made it difficult to tape some interviews.

Students were eager to be interviewed. Some insisted that I could use their names. When I explained the importance of their anonymity to the anonymity of other participants, they agreed to using a pseudonym. Unlike adult participants who engaged in discussion, students waited for me to ask questions from the interview guide. Some students answered questions with brief comments. Prodding often failed to solicit further information. Other students told stories and shared analogies based upon their experiences.

Contributing reasons for variations in responses consisted of students' facilities with the English language, their perceptions of me (trust, mutual respect, and acceptance of their ideas), self-
confidence, level of comfort with being interviewed and taped, personality (extrovert, risk-taking, shy, withdrawn), physical well-being (rested, fatigue, physical wellness, unwell), emotional well-being (stable and unstable relationships with peers, teachers, and family), and academic achievement (impact of low, average, and high marks and grades on self-image).

For example, during one interview I encountered great difficulty eliciting information from a student. Her lack of enthusiasm seemed to be out-of-character in terms of my observations of her classroom involvement. When I asked if something was wrong, she replied that she was very tired. She had been to a rock concert the night before and had been up late. She wanted to finish the interview so we made arrangements to talk at another time.

I explored students' interpretations of classroom assessment-evaluation practices, and their reasons for holding such views. Students described how and why they would change certain practices. Equally important were students' interpretations of how marks, structured written comments, and letter grades affected their
intra- and inter-personal relationships.

Teachers associated with the grade 5/6 class interpreted policy regarding assessment and evaluation differently. I explored the degree to which their interpretations were similar, and ways in which they differed.

I interviewed the principal because I wanted to understand how Grant interpreted the policy. He described how he interpreted policy within the context of the school, and with specific reference to the intermediate division.

Field-Notes. I kept substantive field-notes throughout the course of the study that described daily accounts of circumstances, events, and discussions within which I participated. Informal and formal observations made up part of these notes. Included in substantive notes were my ongoing reflections concerning the conduct of the study in terms of problems, successes, and questions. My methodological notes tracked changes in research questions, data-collection techniques, and the time-line for the study.

I wrote analytic notes directly across from substantive and methodological notes in the field-
diaries. Analytic notes consisted of my reactions to daily experiences in the form of questions, concerns, and hypotheses. By logging my personal reflections, I monitored my interpretations of events and relationships. In other words, I kept track of what pushed my buttons to perceive situations as I did. This was also a useful approach to identify problems before they became serious.

These types of field-notes provided me with a descriptive data-base of classroom practices within and through time. They provided me with evidence of the temporality of Anne's classroom practice. Using content analysis, I analyzed notes and constructed themes regarding issues associated with assessment-evaluation. I compared field-note data with data from documents, observations, interviews, and literature.

Field-notes totalled one hundred and six entries. Quotations and references made to field-notes in the thesis are referred to by FN, followed by the date of the entry.

Data-Analysis

Data-analysis was ongoing throughout the data-collection period and the writing of the thesis. I
analyzed documents, interview transcripts, observations, and field-notes using constant comparative analysis and methodological triangulation.

During the first two readings, I colour-coded data in terms of recurring words, phrases, and images. As the incidence of repetition increased I classified themes according to colour. For example, the colour black signified tests, brown represented portfolios, orange stood for projects, red indicated achievement, yellow signified criterion-referenced evaluation, and so on.

By the second reading I had marked some selections with two or more colours. Layers of colour indicated interconnected themes. I wrote brief notes about the substantive nature of interconnections in a journal. I also recorded themes in their respective colours. Beside each theme I wrote its abbreviated form in pencil. For example, criterion-referenced evaluation became CRE. During the third reading I used abbreviated codes to indicate themes.

When I completed the third reading of data, I returned to my list and added new themes. Under each theme I made key-word notes to describe its major
components. I compared my journal notes concerning interconnected themes with the list to determine how and where themes interconnected. I collapsed closely connected themes to create one category.

Under each category, I wrote key-word notes about each theme. Then I coded each theme-note to indicate its source and date of acquisition. For example, theme-notes for criterion-referenced evaluation were drawn from documents (Doc), interviews (Inter), formal and informal observations (FO; IO), and field-notes (FN). Multiple sources also represented multiple time periods. Data were triangulated along intersecting axes of method, time, space, and people.

In this way, I used constant comparative analysis and methodological triangulation to corroborate multiple perspectives associated with themes and categories. Equally important, I wrote brief descriptions of themes to facilitate my examination of contradictions generated by multiple perspectives.

Content analysis of documents focused on identifying who wrote them, why they were written, for whom, and their major assumptions concerning interpretations of key terminology, practices, and
implementation issues.

When analyses generated contradictions in relation to data from documents, interviews, observations, and field-notes, I examined contradictions in relation to literature. I also took Anne's interpretations of contradictions into consideration. Quotations and references to documents in the thesis are referred to by Doc, followed by the author or acronym describing the document, and year of publication or solicitation. For example, (Doc-RC, 2.3.95) stands for report card document solicited on February 3, 1995.

Content analysis of field-notes identified recurring words, ideas, images, and themes from observations. I re-examined data from documents, interviews, and field-notes with this information. Quotations or references made from field-notes in this thesis are referred to by FN (Field-Note), followed by the date on which the note was recorded.

Informal and formal observations gave me opportunities to explore how talk and actions of teachers and students regarding assessment-evaluation were represented in classroom practices. I corroborated my notes from observation periods with
data from documents, interviews, and field-notes. I also compared observation data with the literature. Information from observation notes was introduced, where appropriate, into interviews for further elaboration by participants. Quotations and references made from observations in the thesis are referred to by IO (Informal Observation) and FO (Formal Observation), followed by the date on which the observation was conducted.

Through content analysis of interview transcriptions I identified recurring words, ideas, and phrases. I used this information to construct themes associated with assessment-evaluation. I compared these themes with data from documents, observations, and field-notes, and literature. Quotations and references made from interviews in this thesis are referred to by Inter (Interview), followed by the date on which the interview was conducted.

Participants were invited to review their transcripts to add, delete, and clarify their statements. Parents did not request transcripts for review. Students read their transcripts but did not make changes. Anne made grammatical changes to her
transcripts. Participants selected and/or were assigned a pseudonym.

Transferability, Reliability, and Validity

I have chosen the criteria of transferability as a measure of the study's usefulness to other researchers (Lincoln & Guba, 1985). Whether or not the significance of the study is transferable to another setting is a question that must be determined by those who are interested in contemplating such a possibility. The degree of transferability will be "...dependent on the degree of similarity between the sending and receiving contexts" (Lincoln & Guba, 1985, p. 297).

In terms of reliability, I have provided a detailed account of the conduct and context of the study so that others can use my work in coordination with their own research interests.

Silverman (1993) explains that triangulation and respondent validation are suitable forms of validation for the logic of qualitative research. I triangulated various types of data to corroborate participants' interpretations across different settings and time periods. I also triangulated multiple data sources to explore relationships among policy and classroom
assessment—evaluation practices.

At regular intervals during interviews I checked my understanding of participants' interpretations by restating their remarks, and asking if my restatements were accurate reflections of their comments. When participants agreed with my restatements, interviews continued. When participants disagreed, or were uncomfortable with my interpretations, we negotiated meaning until my comments met with their approval. 91% of these checks and negotiations were taped.

Anne not only checked her interview transcriptions, but she also read and critiqued each draft of chapter five; the chapter focusing on her classroom practice. Anne edited chapter five by correcting inaccuracies concerning her background, correcting grammatical errors in her transcripts, pointing out claims that needed clarification, and by challenging my use of language when it did not accurately represent her views.

All participants in this study will receive a summary of the thesis. That is, I will send one summary to each family, each teacher, the principal, and the School Board.
I extend my discussion of validity to include the usefulness of the study. Participants found the study useful for several reasons. Grant pointed out that the study was useful because Anne and I worked collaboratively in her classroom. He added that the study created avenues for parents, teachers, students, and administration to share their interpretations and concerns regarding policy and classroom practices.

Anne said the study was useful because it gave her an opportunity to understand concerns of students and parents, and to modify her practice where appropriate. For example, when I explained to Anne that some students were not comfortable with student-led conferences, she decided to modify her approach for the following year. On another occasion, I pointed out that some students thought that self-evaluation was a waste of time because their marks never changed. During the school year 1995-96, Anne made marks negotiable in relation to some activities.

Anne said that she liked to dialogue with me about assessment-evaluation issues because we worked collaboratively. In addition, informally-structured interviews gave her opportunities to sort through her
assumptions, ideas, and interpretations.

Students also liked having opportunities to talk about their opinions and concerns. For some students it may have been the first time that their views had been taken seriously by an adult. Other students had dialogued with parents and adults from a very young age. One girl in grade 6 told me that she liked being able to talk about things that bothered her without fear of recrimination.

Parents appreciated discussing their opinions and concerns during interviews. I invited them to telephone me if they wanted to review their transcripts.

Teachers enjoyed sharing their views and concerns. The logistics of teaching often means that teachers are isolated in classrooms. Rarely are their opinions sought out. Teachers were invited to read their transcripts.

The criteria of dependability for the study recognizes that since reality is socially and culturally constructed, the conduct of the research and the nature of the research design will change in response to contextual changes (Lincoln & Guba, 1985).
Careful and consistent documentation of the research process, the triangulation of methods, and constant comparison of the data, have increased confidence in the significance of these changes and their relationship to the research questions.

Overall, I have learned that extensive collaboration with a teacher and her students allows for an analysis of assessment-evaluation in a context that is useful to teachers.
CHAPTER 4

Return to no-nonsense education lauded: B.C. schools will retain the best of (Year 2000) program — and trash the rest. (Vancouver Sun, 1993, p.B1)

EDUCATION REFORM IN BRITISH COLUMBIA

In this chapter I try to locate policy changes towards classroom assessment-evaluation practices from 1987 to 1997 with particular attention to criterion-referenced evaluation and letter grades. These dates focus debate about assessment-evaluation within a ten-year period. In addition, I examine how the school principal at Westheights understood the policy change.

July 1988, the commission collected approximately 2,350 oral and written public submissions regarding education in BC. Submissions focused on peoples' interpretations of education and the role of schooling in the educational process. In addition, difficulties associated with the provision of educational services in the province were explored.

The Commission recommended that preparation of students for the workplace was an overarching goal of a common curriculum from grades one to ten. It advised restructuring of the "Provincial Learning Assessment Program" to facilitate collection of annual information about student achievement in Humanities, Fine Arts, Sciences, and Practical Arts. Grading systems were to be removed from the primary curriculum to enhance flexibility. Assessments and evaluations of student achievement were to be criterion-referenced and drawn from multiple sources of information (British Columbia Ministry of Education, 1988; 1989a; 1989b).

The Commission also recommended that districts would have greater success obtaining funding in a competitive market by demonstrating increased "value for money" or by improving accountability (Sullivan,
1988, p.49). It suggested that the Ministry of Education, school districts, and schools should work collaboratively to increase accountability. In particular, accountability might be enhanced by developing outcome measures of demonstrated performance that were linked to educational objectives, by interfacing assessment information with data concerning finances and resource management, and by extending improved methods for communicating information to the public.

In 1988, the Royal Commission submitted its findings in the report, "A Legacy for Learners" (Sullivan, 1988). The government adopted seventy-eight of the report's eighty-three recommendations. A fund of $1.4 billion dollars was designated to facilitate implementation of recommendations over a ten-year period.

The "Year 2000." In the fall of 1989, the government issued policy directions for curriculum reform in the document, "Year 2000: A Curriculum and Assessment Framework for the Future" (British Columbia Ministry of Education, 1989a). By May 1990, the document was published as the "Year 2000: A Framework
for Learning" (British Columbia Ministry of Education, 1989b).

The mission statement of the "Year 2000" reflected the Royal Commission's emphasis on the relationship between schooling and work.

The purpose of the British Columbia school system is to enable learners to develop their individual potential and to acquire the knowledge, skills, and attitudes needed to contribute to a healthy society and a prosperous and sustainable economy. (British Columbia Ministry of Education, 1989b, p. ii)

Highlights of the text were learner-focused curriculum and assessment, and continuous progress.

Learner-focused curriculum and assessment is developmentally appropriate, allows for continuous learning....Developmentally appropriate curriculum and assessment take into account the developmental levels of learners at various ages....Assessment procedures concentrate on collecting information about what individual learners can do in relation to standards and expectations about the types of learning that children will exhibit as they grow....With a continuous learning approach, a child is not "failed"....(p. 9)

The document recommended that a criterion-referenced approach to assessment was a component of learner-focused curriculum and assessment (British Columbia Ministry of Education, 1989). Within this context, achievement was appraised according to how effectively a student demonstrated understanding of
curriculum criteria; criteria associated with learning outcomes. Teachers were advised to keep portfolios of work samples that represented student understanding of key concepts and skills.

Teachers will be encouraged to adjust these procedures (informal and formal assessment techniques) to be as learner-focused and criterion-referenced as possible. Teachers will be able to use the learning descriptors provided in Graduation Program documents to help them interpret their observations of student's learning in terms of expectations for learning development. It is expected that for each student a collection of sample work will be maintained, to include especially any work which demonstrates the student's attainment of important learning objectives. (British Columbia Ministry of Education, 1989b, p.27)

"Year 2000" documents drew a distinction between the terms, assessment and evaluation (British Columbia Ministry of Education, 1989a; 1989b). Assessment represented systematic procedures for collecting information concerning what a student was able to do in terms of the "Goals of Education." Evaluation represented ongoing procedures for determining appropriate programs for students based upon assessment information.

Teachers were encouraged to use multiple methods to collect assessment information from multiple sources within the context of a "learner's production of
something, eg. a piece of writing, an oral presentation, a drawing, a constructed object" (British Columbia Ministry of Education, 1989b, p.1). They were advised that ongoing and readily available assessment methods enabled students to demonstrate what they were able to do at different points in their learning. In addition, students needed to understand learning outcomes for assessment procedures to be meaningful, and self-assessments and self-evaluations engaged students directly in their learning.

"Year 2000" policy discussions pointed out that informal and formal reports of student progress should be conducted at regular intervals throughout the school year (British Columbia Ministry of Education, 1989b). In addition, reports should describe students' attainments and emphasize what they were able to do in relation to expected learning outcomes. Teachers were encouraged not to compare performances among students to generate descriptions of individual performance. If parents asked how a child's performance compared to the performance of peers, teachers were advised to reframe the discussion in terms of provincial standards for development and learning. The Draft document, "The
Intermediate Program: Foundations" (British Columbia Ministry of Education, 1992a) explained that "Reference Sets" were being developed to provide teachers with general patterns of development and achievement within particular contexts for learning.

"Year 2000" policy discussions indicated that student performance should be reported via anecdotal report cards, teacher-parent conferences, and student-led conferences (British Columbia Ministry of Education, 1989a; 1989b). Although letter grades were removed from the primary division, their use with intermediate students was optional. The support document, "Intermediate Program: Foundations," provided a compelling discussion of drawbacks associated with the sole use of symbols to describe student achievement (British Columbia Ministry of Education, 1992a). I interpreted the document's use of the term, symbols, to involve letters, numbers, and check marks.

While a concise means of communicating student learning is needed, symbols reduce the complexity of the accomplishments of learners, and oversimplification occurs. The symbols may be given overweighted significance; parents/guardians and students may accept them as unchangeable and unquestionable and rate students' abilities accordingly. Using symbols to compare students
has the potential for labelling, which may lead to negative effects on learner self-esteem, reduce motivation, and seriously impede achievement. While the pursuit of a higher symbol may motivate a student, the symbol often becomes the goal, not the learning. The true value of learning becomes extraneous and is lost as attainment of the symbol becomes central. (British Columbia Ministry of Education, 1992a, p.118)

The text also advised that learner-focused symbols must be criterion-referenced. Moreover, procedures for constructing symbols should include student participation in negotiating and discussing evaluation criteria.

To the extent that symbols can be learner-focused, criterion-referencing is mandated. It is the process of arriving at the symbol that is important, involving negotiation and discussion of criteria with students....they should accordingly be based upon such things as predetermined criteria and information derived from reference sets and student work. (British Columbia Ministry of Education, 1992a, p.118)

Within two years, the NDP government rearticulated the "Year 2000's" emphasis on criterion-referencing by mandating "criterion-referenced letter grades" from grades 4 to 12, and encouraging teachers to use "criterion-referenced evaluation" in their classroom practices (British Columbia Ministry of Education, 1989a, 1989b, 1994e, 1994i). The policy change appeared to turn on uses of terms, criterion-
referenced and evaluation. Whereas "Year 2000" policy discussions linked the term, criterion-referenced, to assessment and evaluation symbols, the "K-12 Education Plan" linked the term, criterion-referenced, to evaluation and letter grades.

The New Democrats take on assessment and evaluation. Allegations of corruption against the Vander Zalm government and increasing support for the Liberal Party ended the reign of the Social Credit government in 1991 (Crawley, 1995). However, the NDP did not come to power in British Columbia because the majority of voters had swung left-of-center. The Liberal and Social Credit parties secured 57.3 percent of the popular vote. The NDP party acquired only 40.7 percent of the popular vote, but gained a majority of seats (Crawley, 1995).

After two years of efforts by the new government to implement the "Year 2000" (1989b), Premier Harcourt stunned the province with his announcement that "...the report card on Year 2000 is in and it's failed the grade" (Vancouver Sun, 9.4.93. p.A1). He based his decision to discard the Year 2000 on the results of a public opinion poll that reflected widespread
dissatisfaction with the school system (Crawley, 1995). The poll's findings indicated that members of the public wanted to improve the quality of education by returning to basic skills and more specific standards. On September 15, 1993 Premier Harcourt appointed Art Charbonneau as Minister of Education.

In a news release on July 8, 1994 the Ministry of Education announced that grades 4-12 students "...will receive mandatory letter grades...so parents have a better understanding of how their child is progressing.... Along with the traditional letter grades of A, B, C+, C, C-, and, F, the IP, or In Progress, has also been added" (British Columbia Ministry of Education, NR29-94, p. 23). In addition, structured written comments would accompany letter grades for grades K-7 students.

Mr. Charbonneau, Minister of Education, explained that changes to reporting procedures responded to parents' concerns.

Children have been bringing home report cards that their parents feel are not clear enough....We are providing clear guidelines which will help teachers and parents evaluate a student's strengths and weaknesses. These changes are fundamental elements of the education reform currently taking place in B.C. schools. (British Columbia Ministry of Education, NR29-94, p. 23)
Furthermore, parents' concerns over anecdotal reporting procedures motivated the proposed reforms.

The new reporting standards replace anecdotal reports. The change is part of the improvements to B.C.'s education system announced last November by Premier Mike Harcourt and Charbonneau. Charbonneau said the new policy is in response to parents' concerns over the anecdotal reporting system, and school boards must adapt their reporting policies to comply with the Ministry's guidelines. The reporting standards have been finalized following consultation with parents and teachers. (British Columbia Ministry of Education, NR29-94, p. 23)

In September 1994, Premier Harcourt and Mr. Charbonneau released the, "Kindergarten to Grade 12 Education Plan" (British Columbia Ministry of Education, 1994i). Policy documents, "Report to Parents" and "Parents' Guide to Standards," were released to the general public (British Columbia Ministry of Education, 1994f; 1994h). Additional documents described policies and implementation plans for members of the educational community (1994b).

Documents regarding reporting procedures for the school year 1994-95 were released to the executive committee of the Ministry of Education, the Provincial Education Committee, the school Superintendents and School Board Chairpersons. Documents included a "Student Progress Report Order" (Ministerial Order,

The "Student Progress Report Order" stipulated that grades 4 to 7 progress reports must, in relation to expected learning outcomes, would include letter grades unless a Board designated an alternative approach for reporting this information to parents (British Columbia Ministry of Education, M191/94). Reports must also include structured written comments concerning what a student was able to do, areas needing extra attention, and approaches for supporting the student's education in relation to expected learning outcomes. Students identified as having special needs were to receive written reports describing their progress in relation to the expectations of their Individual Education Program.

The "Provincial Letter Grades Order" identified letter grades as A, B, C+, C, C-, IP (In Progress), F (Failed or Failing) W (Withdrawal), SG (Standing
Granted), and, TS (Transfer Standing) for grades 4 through 12 (British Columbia Ministry of Education, M192/94). Letter grades ranked a student's performance from outstanding or excellent to unacceptable in relation to expected learning outcomes. Percentages accompanied letter grades A through C-, for grades 11 and 12 courses. The "Guidelines For Student Reporting" explained that letter grades were to be "criterion-referenced" (British Columbia Ministry of Education, 1994e). Structured written comments on grades 4 to 7 intermediate report cards were to include descriptions of behaviour, attitude, work habits, and effort.

Three formal and two informal progress reports were required per student per year. Informal reporting procedures involved teacher-parent and teacher-student-parent conferences, telephone calls and, written or verbal interim reports. A minimum of one teacher-parent conference per student was required (British Columbia Ministry of Education, 1994e).

The symbol, "IP" (In Progress), was to be used when a student needed more time to meet expected learning outcomes. "In Progress" resonated with the "Year 2000" focus on "continuous progress" (British
Columbia Ministry of Education, 1989b). IP on a student progress report was accompanied by comments that described areas needing support, and/or certain requirements that needed to be met. Teachers were instructed to outline the time period that a student needed to meet the expected learning outcomes, and the time needed to review and evaluate a student's performance.

Parents of students in grades 4 to 7 were to be consulted when IP was used on successive reports. When a final report included an IP, information was to be made available to the next teacher. This information involved descriptions of a student's needs and ways her or his program could be supported. The designation, IP, was to be transformed to a letter grade on the first report card of the following school year, and/or when a student's records were forwarded to another school. Use of IP on progress reports was made mandatory for the school year 1995-96 (British Columbia Ministry of Education, 1994e).

Special needs students who were expected to meet the learning outcomes of the provincial curriculum received letter grades and structured written comments
on report cards. Letter grades were regarded as inappropriate for students who failed to meet these standards. Structured written comments replaced letter grades to describe student progress.

The Ministry of Education provided school boards with several planning guides to support implementation of reporting policy. "The Kindergarten to Grade 12 Education Plan" outlined the ways the new reporting policy differed from previous requirements in the following manner:

~ Grades 4 to 7 teachers are required to report using letter grades as well as written comments.
~ The descriptions of the letter grades are revised.
~ "I" for incomplete is replaced by "IP" for "in progress".
~ There are some new guidelines for the use of "F" for fail.
~ Letter grades "D" and "E" will no longer be used.

(British Columbia Ministry of Education, 1994i p. 15).

Included in the document was a description of letter grades. Letter grades symbolized how effectively a student demonstrated understanding of curriculum criteria in relation to provincially
prescribed learning outcomes.

The "Guidelines for Student Reporting for the "Kindergarten to Grade 12 Education Plan" described assessment, evaluation, and reporting in the following ways (British Columbia Ministry of Education, 1994e):

- Provincial Curriculum Guides define expected learning outcomes that are intended to guide the setting of curriculum criteria.
- Assessment and evaluation informs instruction. The Ministry Resources intended to support the practices of assessment and evaluation are: the "Writing and Reading Reference Sets" (British Columbia Ministry of Education, 1992e; 1994c); "Supporting Learning: Understanding and Assessing the Progress of Children in the Primary Program A Resource for Parents and Teachers" (British Columbia Ministry of Education, 1992b); and, the "Assessment Handbook Series: Student-Centred Conferences, Portfolio Assessment, Student Self-Assessment, Performance Assessment" (British Columbia Ministry of Education, 1994a).
- Reporting policy is supported by the "Guidelines for Student Reporting" (British Columbia Ministry of Education, 1994e). Informal and formal reporting
procedures set goals for student learning.

Following the text of the "Year 2000 - A Framework for Learning" (British Columbia Ministry of Education, 1989b), the "Guidelines for Student Reporting" and the "Assessment Handbook Series" (British Columbia Ministry of Education, 1994e; 1994a) drew distinctions between assessment and evaluation. NDP policy documents described assessment as systematic procedures for collecting information about what students knew, were able to do, and tried to do. Evaluation was described as procedures for making judgements and decisions about student performance based upon interpretation of assessment data. The new guidelines expanded the "Year 2000's" interpretation of assessment, as a process for collecting information about what students were able to do, to include what students knew and tried to do. Both Social Credit and NDP assessment policies recommended that assessment and evaluation in classrooms compare student performance with curriculum criteria or provincially prescribed learning outcomes. "Expected learning outcomes" were provincially prescribed indicators of what students were expected to learn (British Columbia Ministry of Education, 1994e).
Following the "Year 2000's" focus on "expectations for learning development," documents associated with the "K to 12 Education Plan" focused on "widely held expectations" (British Columbia Ministry of Education, 1989b; 1989b; 1992b; 1994a; 1994e; 1994i).

Following the "Year 2000," the "K to 12 Education Plan" expanded reference sets to assist teachers in setting performance standards. These sets included, "Evaluating Writing Across Curriculum" and "Evaluating Writing Reading Curriculum" (British Columbia Ministry of Education, 1992c; 1994c).

In June 1995, the Ministry distributed three "Integrated Resource Packages" (IRPs) to schools across the province (British Columbia Ministry of Education, 1995a; 1995b; 1995c). IRP's outlined prescribed learning outcomes and suggested instructional techniques, assessment strategies, and resources for science, mathematics, and personal planning from kindergarten to grade 7. Implementation of IRP's was scheduled for September 1995. The preface to the "Science K to 7 Integrated Resource Package" described relationships among key concepts in the following manner:
The learning outcome statements are content standards for the provincial education system. The learning outcomes are statements of what students are expected to know and do at an indicated grade; they comprise the prescribed curriculum. Learning outcomes are clearly stated and expressed in measurable terms. The outcomes are benchmarks that will permit the use of criterion-referenced performance standards. Student performance will vary in relation to the outcomes. Evaluation, reporting, and student placement with respect to these outcomes is dependent on the professional judgement of teachers, guided by provincial policy. (British Columbia Ministry of Education, 1995c, p. III)

The "Guidelines for Student Reporting" (British Columbia Ministry of Education, 1994e) explained that teachers' assessment procedures involved "...observation, student self-assessment, daily practice assignments, quizzes, samples of student work, pencil-and-paper tests, holistic rating scales, projects, oral and written reports; reviews of performance, and portfolio assessments" (p. 14). Teachers' evaluations of student progress were based upon multiple forms of assessment data.

The "Guidelines for Student Reporting" described "criterion-referenced evaluation" as an approach that compared student performance to curriculum criteria and was regarded as suitable for classroom use (British Columbia Ministry of Education, 1994e). In contrast,
the document stated that "norm-referenced evaluation" compared student performance with performance of other students and was regarded as suitable for large-scale system assessments.

Norm-referenced evaluation....is based on a 'normal distribution.' A normal distribution shows how achievement in a particular area is distributed over an entire population. To use norm-referenced evaluation appropriately, a student's achievement must be compared with a reference group large enough to represent the population. Placing student achievement on a curve does not accurately describe a student's individual progress; it compares student achievement to that of others rather than comparing how well a student meets the criteria of a specified set of learning outcomes. (British Columbia Ministry of Education, 1994e, p. 16)

Even though norm-referenced instruments were regarded as unsuitable for use by classroom teachers, candidates for financial awards and Special Education support were in part identified by means of norm-referenced procedures.

The "Guidelines for Student Reporting" described steps involved in a teacher's use of criterion-referenced evaluation in the classroom. A similar outline was provided by the IRP's (British Columbia Ministry of Education, 1994e; 1994a).

1. Identify the expected learning outcomes (as stated in curriculum guides).
2. Identify the key learning objectives for students.
3. Establish and set criteria.
4. Involve students, when appropriate, in establishing criteria.
5. Plan learning activities that will help students gain the knowledge or skills outlined in the criteria.
6. Inform students about the criteria their work will be evaluated against prior to the learning activity.
7. Provide examples of the desired levels of performance.
8. Implement the learning activities.
9. Use various assessment methods based on the particular assignment and students.
10. Review assessment data and evaluate each student's level of performance or quality of work in relation to the criteria.

Although the document recommended that negotiation of criteria should occur when it was "appropriate," no criteria were provided to help teachers decide what counted as appropriate. In addition, assessment methods and related data were important components of criterion-referenced evaluation.

Was criterion-referenced evaluation new? Yes and no. Both Social Credit and NDP assessment policy recommended that evaluation techniques and procedures be criterion-referenced. Both assessment policies interpreted criterion-referenced approaches to assessment and evaluation in terms of techniques that
compared student progress with curriculum criteria or learning outcomes. Within this context, criterion-referenced evaluation was not new; keeping in mind that the "Year 2000" did not use the terminology, criterion-referenced evaluation (British Columbia Ministry of Education, 1989a; 1989b; 1994e). Criterion-referenced evaluation was legislated by the "Year 2000" on the basis of intent; not on the basis of terminology. Unlike the "K to 12 Education Plan" that described criterion-referenced evaluation as a series of steps for implementation by teachers, the "Year 2000" discussed criterion-referencing as an approach to assessment and evaluation.

The Standpoint of a Principal

Grant, the principal of Westheights school, explained that the policy change was new to everyone, and that everyone had to work together to figure out what it meant (FN, 06.12.94). He added that even though he was a principal, he did not have the answers to how policy should be interpreted.

Grant recommended that I attend a series of Professional Development Day meetings with staff regarding letter grades and criterion-referenced
evaluation (FN, 06.12.94). During the meeting on criterion-referenced evaluation, one intermediate teacher asked "Jeff," the District Principal, what motivated the new reporting policy. Jeff replied that,

Letter grades are intrinsically political and...anecdotal comments were translated to structured written reports to placate the public as a measure of accountability. (FO, 31.01.95, p. 10).

Grant responded that he had not received complaints from Westheights parents about the school's reporting practices (FO, 31.01.95, p. 10). He pointed out to Jeff that if new reporting policy was a response to public dissatisfaction over reporting practices, then he did not understand why Westheights parents were silent on the matter. Several teachers commented that they also had received no negative feedback from parents concerning reporting practices. Jeff added that, although parents as a group tended to say that "public education is going down the tubes," they appeared to be happy with their local school (FO, 31.01.95, p. 11).

Grant and Jeff expressed agreement that parents and the public in general placed a great deal of importance on letter grades. Grant thought that
Ministry policy regarding letter grades from grades 4 to 7 was created "...in response to parents wanting something specific that they could attach a meaning to" (Inter, 26.05.95, p. 7).

Grant: ...my sense after 30 years is that for most parents what they still want to know is how is my child doing compared to the rest of the class....(p. 12)

Jeff said that symbols like letter grades had the power to arbitrate self-esteem and group membership within Canadian culture.

The power of the symbols regarding inclusion and self-esteem in our culture is amazing. (FO, 31.01.95, p. 13).

After the meeting, Grant told me how badly he felt after witnessing the degree of resistance his staff exhibited towards criterion-referenced evaluation and letter grades. As principal, he was responsible for ensuring that his staff complied with provincial and school board assessment policy (Reporting Student Progress, British Columbia Ministry of Education, September 1994).

Grant: ...you sat in an afternoon session with the Board resource people and you saw the body language around the room. You heard the lack of questions, which could be interpreted as complete lack of interest in what we were trying to do....I felt really embarrassed that afternoon. I thought I've done what I can to try to raise awareness
The resource person (Jeff) said to me as he's walking down the hall afterwards, 'I think I blew it.' I said, 'No, I don't think you blew it, you did the best you could. You know I can't be accountable for how 35 or 40 people will react, I can understand why you might have felt that you blew it'. (Inter, 26.05.95, p. 24)

Grant went on to describe difficulties school administrators encountered in translating Ministry and Board policy into "action" (Inter, 95.05.26).

We are school-based management and our role is to make sure that the policy that the Ministry and Board is establishing is implemented. Now...how does that translate into action?...first of all we have to try and be sure that we understand what that policy is so that...we're clear on what it is we're being asked to do. And if that means we need to undertake some reading or some inservice ourselves...to be able to...be clear that we understand not only what the policy says but what it actually means...that's probably the first step....the system immediately starts to break down because there is no consistent mandatory inservice for administrators to make sure that in fact everyone does understand....in the same way teachers are not forced to attend professional development on a particular topic....(Inter, 26.05.95, p. 23)

Grant interpreted the Ministry's emphasis on criterion-referenced letter grades as the most critical policy change for teachers (FN, 06.12.94, p. 151). He thought that the shift from anecdotal to structured written comments was less critical (Inter, 18.05.95, p. 1). Structured written comments differed from anecdotal comments in that the former were more
specific (Inter, 26.05.95, pp.1-2). Although the shift to structured written comments was not a major change for primary teachers, it appeared to be significant for intermediate teachers.

Grant: ...it's a bit of a shift from just...the anecdotal notion where sometimes people rambled on at length about what kids could do, but weren't necessarily as specific about areas needing improvement or work, and certainly the expectation that there be a clear statement of goals for the next term, which could include what the teacher's going to do, and what the parent could do to support that kind of clarifying purpose, of goals....So for the intermediate that is a change because what the intermediate report card consisted of before were the letter grades and work habit marks and a space for some anecdotal comment or elaboration, but there was never the direction in terms of the structured written report.... (Inter, 26.05.95, 1-2)

Grant added that some intermediate teachers were disturbed by the amount of writing required to create structured written comments for each subject area on the report card.

And of course it caused a lot of anxiety ...because it looked as if it was going to require a tremendous amount more than teachers had been used to having to write....the question was raised is that going to be at each report card time, because if there's five or six or seven subjects involved, and if kids are platooning to various specialists, and each of those specialists is expected to contribute a written comment about their subject, this is a huge additional amount of work for the intermediate teachers beyond what they've been asked to do. (Inter, 26.05.95, p. 2)
Grant assumed that teachers across the province shared these concerns. He explained that the Ministry of Education responded to the problem by stating that, although comments must be written for each subject area on the report card, comments need only be written once a year (Inter, 26.05.95, p.3).

Grant decided that the Ministry intended that teachers write comments on report cards for Language Arts and Mathematics for each reporting period. The remaining subjects could be addressed once during the year. Jeff advised staff that structured written comments included comments on attitudes, work habits, and effort (Formal Observation [FO], 31.01.95, p.9).

Grant added that the shift to criterion-referenced evaluation was a separate issue from the shift to structured written comments. He pointed out that intermediate teachers on his staff had always used letter grades (FN, 06.12.94, p. 151). However, teachers differed in their methods of constructing letter grades.

Grant: ...there's nobody in that intermediate staff here right now who isn't used to giving letter grades, and how they've been deriving those letter grades probably is in a variety of ways. But more accurately they probably have been norm-referenced. I see different people, they add up
their marks, they change them into percentages, whether they put them on a bell curve or what they do. I'm not necessarily even clear how everyone has done it, but through years of experience and practice, everyone has their way of coming up with what that means. The ways they have in the past derived those letter grades and some, it's been the very traditional old-fashioned way of saying 86% and up is an A, and 78 to 85 is a B or whatever those numbers are. And for a lot of people it's just been a case of adding up a whole bunch of marks and literally dividing them into a percentage and then assigning some letter grade. The shift to devising a system that is no longer norm-referenced but criterion-referenced I think is a major shift for many of these people. (Inter, 18.05.95, pp. 1-2)

Grant explored the shift from norm-referenced to criterion-referenced evaluation through his deliberations concerning the meaning of norm-referenced.

Grant: ...even the notion of norm-referenced assessment is...open to all kinds of interpretation because in the strictest sense of the word, if we're talking about norm-referenced we usually are talking about something that has developed a series of norms based on a fairly large population, not even, you know even a classroom population of 30 or 36, 30 today, would under the original definition of norm referencing be a large enough population....(Inter, 18.05.95, p.3)

...in actual fact what we've been doing for years may not even be truly norm-referenced because we have been evaluating kids based largely on teacher's designed tests or assignments which may be totally different from what we gave the previous year...and so how we can even say that we have been relating them to a norm is probably open for all kinds of debate anyway. (Inter, 18.05.95, p. 4)
He added that instructional strategies associated with criterion-referenced evaluation opened the possibility for all students in a class to get an A in a particular assignment.

Grant: ...if you have collectively established a group of criteria and you make sure that the children understand them and there's been some input and some discussion around that, it's conceivable that...everyone could get an A, on any given task, product, process, whatever. And that would be a huge difference from the notion that there will be some As and some Bs and some Cs and god forbid some Ds and some Es. (Inter, 18.05.95, p.6)

Although he spoke positively about criterion-referenced grading practices, Grant was irritated by the Ministry's claim that letter grades for intermediate students may, or may not, be reported via the report card. He pointed out that if teachers were not required to write the letter grade on the report card, then "what were we going to have to ensure that in fact the letter grade had been assigned, reported, and was documented?" (Inter, 26.05.95, p.7). He added that when 75 Boards had the option to develop local policy concerning the communication of letter grades to parents, the potential for confusion over student achievement was exacerbated by students transferring from one district to another.
Grant explained that teachers, students, and parents would not embrace the policy change until they understood what it meant, and that common understanding needed to begin with teachers. Furthermore, different groups of people have different stakes in seeking understanding (Inter, 18.05.95).

...do they have the understanding?...I don't know if we can talk about all of the people in the same way because we're really talking here about kids, parents and teachers. (p. 6)

Grant pointed out that some people and institutions placed a high value on letter grades as exchange for money and educational opportunity.

...I think we have a lot of evidence in this school that there are still a lot of people out there for whom the ultimate importance of what the kids do is the letter grade on the report card.

...when kids are still motivated by, they get $20.00 for every A...where's that coming from?....And unfortunately part of that is driven by other institutions outside of our domain which are still looking for As and Bs or 86%...(Inter, 18.5.95, p.8)

Conclusion

Some observers interpreted changes to educational policy by the NDP government as thinly veiled efforts to "...look decisive and...to put its own stamp on an education program designed by the previous party in power" (Crawley, 1995, p.7). My analysis of policy
documents from 1987 to 1997 indicated that key aspects of "Year 2000" approach to assessment and evaluation were present in the "K to 12 Education Plan" (British Columbia Ministry of Education, 1989a; 1989b; 1994e).

The "Year 2000" and the "K to 12 Education Plan" shared a concern for clearer standards of achievement and recommended multidimensional and criterion-referenced approaches to assessment and evaluation. Both governments recommended that teachers involve student participation in setting performance criteria where appropriate. They also promoted performance assessment, portfolio assessment, student self-evaluations/self-assessment, peer-evaluation, outcomes-based education, reference sets, and Provincial Learning Assessments.

I detected differences between the Social Credit's and NDP's approaches to assessment and evaluation. Whereas "Year 2000" policy discussions used the term, criterion-referenced, within the context of approaches to assessment and evaluation, the "K to 12 Plan" used the term, criterion-referenced, in relation to evaluation and letter grades (British Columbia Ministry of Education, 1989a; 1989b; 1994e).
The "Year 2000" used the term, criterion-referenced, to refer to approaches to assessment and evaluation that were referenced to curriculum criteria and learning outcomes. The "K to 12 Education Plan" specified steps for teachers to follow when using criterion-referenced evaluation in classrooms. Finally, whereas "Year 2000" policy recommended that letter grades be used where appropriate, the NDP government mandated criterion-referenced letter grades from grades 4 to 7.

The NDP's use of criterion-referenced in relation to evaluation and letter grades was confusing. Policy documents described criterion-referenced evaluation as an approach wherein teachers ensured that students understood, and even set, performance criteria before beginning a task. Criterion-referenced letter grades were described as classifications of student performance compared with expected learning outcomes.

What did criterion-referenced mean within the context of criterion-referenced evaluation? Why was the term not used in relation to assessment? When did criterion-referenced evaluation involve assessment methods? When did assessment methods involve
comparisons of individual performance with criteria that students were aware of, and even participated in setting, prior to beginning a task? How were teachers to decide when it was appropriate to involve student participation in setting criteria?

What did criterion-referenced mean within the context of criterion-referenced letter grades? Were criterion-referenced letter grades constructed from data wherein students were aware of curriculum or performance criteria, and when appropriate, they participated in setting criteria? And/or were criterion-referenced letter grades constructed from data wherein individual performance was compared with curriculum criteria and learning outcomes?

It is not surprising that Grant was confused by the policy change. The power of language to mediate confusion is illustrated by difficulties associated with the NDP's uses of criterion-referenced, assessment, and evaluation. Within this context, vague, ambiguous, and even incoherent policy discussions generated bewilderment.

In the next chapter, I examine how Anne tried to make sense of the policy change within the context of
her classroom practice.
CHAPTER 5

I guess I just do whatever to me seems what a person should do if they're a teacher. (Anne, Interview, 96-04-24)

The 'feel' for the game is the imminent sense of the future of the game, the sense of direction of the history of the game that gives the game its sense. (Bourdieu, 1990, p.82)

THE EXPERIENCE OF A TEACHER

In this chapter I examine Anne's understanding of, and responses to, the policy change in her classroom practice. Whereas policy discussions defined assessment as procedures for gathering information about student progress, Anne defined assessment as classroom practices for collecting and making judgements about student learning based upon curriculum criteria. Anne's understanding of assessment paralleled my understanding of assessment-evaluation.

Anne's Background and Philosophy of Education

Anne began her career as an elementary classroom teacher and teacher-librarian. Six years later she took a two-year leave from teaching to raise her first child. She returned to work on a part-time basis until her second son was born. Anne then took a six-month
leave to be with the two children. Following this period, she worked as a teacher-librarian on a part-time basis until her youngest son entered grade 8. At this time, she returned to full-time teaching. Throughout the following years she taught a grade 4 class, two grade 4/5 classes, and two grade 5/6 classes.

Anne explained that her parents' emphasis on family led her to place great importance on her own family, and influenced her to see students as extended family (Inter, 4.28.96). Unity, sharing, respect, and tradition were important values for Anne at home and at school.

My family is really important to me. My parents always stressed family and family traditions that I have tried to carry on in my family because I have felt they were important. It's important to have a sense of family. I suppose if you have a sense of family and think that family's important it perhaps does reflect in your teaching. You try and make your class feel like a family and make everyone feel together, especially if you have a combination class. You don't want to make them feel like each of them is a different grade. You try and make everyone feel as if they do belong to one common group, a family so to speak. I always try to make everybody in the class feel very comfortable, relaxed, and welcome. I encourage everyone to be polite and not put each other down, so that people feel safe and supported. I suppose all this relates to my sense of family. (Inter, 4.28.96)
Presently, Anne's husband is a high school teacher. Her sons are taking college courses and working part-time.

Anne added that university studies influenced her philosophy of education and classroom practice. During her teaching career she took courses leading to a diploma in Special Education. Anne pointed out that courses ranging from learning disabilities to giftedness sensitized her to the diversity of students' needs (Inter, 96-04-28).

...the Special Education courses that I took probably gave me a good understanding of all the different kinds of learners, and the frustrations that they would feel, and the needs that they might have. I think those were very useful courses to have as a background. I guess I like to perhaps use a gifted education model of teaching, wherein activities are open-ended so that all learners can approach the activities and enter at their level....I like to use gifted education approaches like creative problem-solving, future problem-solving, creative thinking skills, and critical thinking skills with all students because I believe that its really important for all students to learn to think and to solve problems, to think critically and creatively....especially the lower-end students because I think sometimes they lack self-esteem and they just think that they can't come up with answers....It's really important for students to make important decisions about their learning, to get to know themselves as learners. (Inter, 4.28.96)

At home, and in the classroom, Anne emphasized the
importance of mutual respect. She saw mutual respect as a unifying code of behaviour embodied in classroom relationships.

I like to try and unite students as soon as possible into a cohesive group where they can feel comfortable and safe with each other, so that they can give oral presentations, and state their opinions freely, and know that they won't be laughed at or criticized. It's important, that I respect them, their opinions and thoughts, and that they respect my opinions and thoughts. (Inter, 4.28.96)

Anne's educational philosophy centered around her assumption that the purpose of education was to empower students to "learn how to learn" (Inter, 4.28.96).

I want students to be independent learners and learn how to learn, take charge of their learning. For me to be their guide, and not the person who supposedly knows all the information and I just give it to them and they learn what it is I tell them to learn. The main thing is, if you don't know something and as long as you know how to find the information, that's the important thing. (Inter, 4.28.96)

Making Sense Of The Policy Change

Anne's understanding of criterion-referenced evaluation was shaped by her family background, teacher education, classroom practice, philosophy of education, colleagues, professional development workshops, professional reading, university course work, and Ministry and District Office assessment and evaluation.
policy. Her power to interpret criterion-referenced evaluation in her classroom practice was limited by her perceived obligation to implement Ministry and School Board assessment and evaluation policy.

Anne first encountered the term, criterion-referenced evaluation, during District Office and school-based workshops (Inter, 5.22.95). On several occasions, workshop leaders used the terms, criterion-referenced assessment and criterion-referenced evaluation, interchangeably (FN, 11.3.94; FN, 1.30.95; FN, 1.31.95). Both expressions were used to refer to practices associated with involving student participation in setting criteria, collecting information about student progress, and making judgements about student progress based upon curriculum criteria (Doc, 11.3.94; Doc, 1.30.95; Doc; 1.31.95; Doc, 2.10.95).

On November 3, 1994, Anne and I attended a District Office professional development workshop on criterion-referenced evaluation. "Sandra Evans," the workshop leader, talked about criterion-referenced assessment throughout the meeting. She did not use the term, criterion-referenced evaluation (FN, 11.3.95).
Participants in the workshop included approximately twenty-eight invited teachers and administrators. Sandra, a consultant from a neighbouring board, began the meeting by asking, "What kind of intellectual capital do we need to build in these children? How do you get and keep a job in the 21st. century?" (FN, 11.3.94, p.1). Without pause she responded, "Where children are involved in process. Where they negotiate criteria with teachers" (FN, 11.3.94, p.3).

Sandra continued by asking, "What is the most important piece for you to get from criterion-referenced assessment?" Participants broke into groups to brainstorm "needs, wants, and wishes" (FN, 11.3.94, p.3). Each group chose one person to record their ideas on chart paper. The recorders read their respective charts to the group, and taped the charts on the walls around the room.

During her talk about criterion-referenced assessment Sandra distributed a series of handouts that described steps for implementing criterion-referenced evaluation in the classroom. The manner in which she introduced this material implied that criterion-
referenced assessment and criterion-referenced evaluation meant the same thing. Moreover, one handout included steps for implementation of criterion-referenced evaluation. Attention was drawn to those steps wherein students were to be involved in establishing criteria when appropriate, and students were to be made aware of evaluation criteria before engaging in an activity (Doc, 11.3c.94).

Sandra then engaged the group in an activity called the "International Clapping Institute" to illustrate the importance of discussion and negotiation when setting criteria with students; discussion and negotiation improved performance. Four people from the audience volunteered to be clappers and five volunteered to be assessors (FN, 11.3.94). Sandra's use of the term assessor as an individual who evaluated clappers conflated the terms assessment and evaluation.

Clappers introduced themselves and described how they became interested in clapping. "Jack," the first clapper, clapped his hands and then was asked by Sandra to leave the room. In Jack's absence Sandra told the assessors to rate his clapping on a scale of 1 to 4. She recorded and totalled the assessors scores on an
overhead projector. Jack was called back into the room and told his score.

Before the second participant clapped, Sandra placed the "international criteria" for clapping on the overhead projector. Criteria involved "volume," "appropriateness" and "creativity." The second clapper was rated by the assessors on a scale of 1 to 4 using the international criteria.

Sandra asked the final clapper to explain where she would be doing her clapping, and the importance of her clapping. After her clapping was rated, Sandra told the woman that she could ask questions of the assessors regarding her evaluation. The woman asked assessors to clarify what they meant by their use of the term, creativity. Then, Sandra asked her how she felt about the criteria of creativity, and what kind of assessment she would like to have.

Sandra's use of the term, assessment, as a kind of evaluation reinforced my belief that classroom procedures and practices associated with assessment and evaluation were interconnected to such a degree that it was very difficult to avoid using the terms interchangeably; hence my preference for the expression
assessments-evaluation.

The last clapper described her criteria for creativity and said that she wanted her assessment to take the form of an unstructured written comment. Sandra ended the activity by asking the remaining two clappers to share how they felt about the assessment process. She pointed out that the activity of the clapping institute illustrated the importance of setting criteria with participants so that they knew how they were assessed before engaging in an activity, to ensure that criteria was appropriate, and to clarify the meaning of criteria (FN, 11.3.94). Although these points were illuminated by the activity, the distinction between assessment and evaluation was unclear.

In the afternoon several teachers gave the following testimonials regarding their experiences with criterion-referenced assessment: "Standards have gone up, atmosphere has changed; Language has changed to criteria-language; Language has changed in the classroom" (FN, 11.3.94). Towards the end of the day, Sandra pointed out that letter grades were constructed from comparisons of student work with classroom
criteria that represented intended learning outcomes of the curriculum.

We're not comparing kids. You have to look at the ends and measure criteria in terms of how well the outcome is reached. So you are measuring the path of the outcome to be achieved and it doesn't have to be the entire outcome. You can compare the child's sample or product to the classroom criteria...compare it to reference sets and national standards. (FN, pp. 59-60)

Sandra ended the workshop by promoting a handbook entitled, "Criterion-referenced Evaluation: Getting Started" (Doc, 11.3b.95). Her closing comments were, "The road to success is always under construction" (FN, 11.3.95, p.79).

Unlike Sandra, Anne did not use the term criterion-referenced assessment. Instead, she adopted the term, criterion-referenced evaluation, used by the Intermediate Coordinator and District Principal in their professional development workshops (FN, 1.30.95; FN, 1.31.95).

Sandra's emphasis on setting criteria with students strongly influenced Anne to interpret criterion-referenced evaluation as, in part, an instructional strategy (Inter, 7.27.95). She explained that Sandra's workshop led her to recall instructional strategies that she was already using in the classroom.
When I first heard about criterion-referenced evaluation, it made a great deal of sense. When Sandra talked about teaching and learning strategies that I have always liked using in the classroom, that was an important connection for me in relation to criterion-referenced evaluation. (Inter, 5.28.95, p. 11)

Anne added that developing criteria with students was a new idea for her.

Developing criteria with students was a new idea for me this September. Maybe other people knew about it but I was unaware of the approach. (Inter, 5.22.95, p.17)

Jeff, a District Principal, conducted a professional development workshop on criterion-referenced evaluation for staff of Westheights' school on January 31, 1995. During the workshop he distributed a handout entitled, "Criterion-referenced Evaluation." Just like the information in Sandra Evans' handouts, Page 6 of Jeff's booklet outlined the steps teachers were to follow to implement criterion-referenced evaluation in their classrooms (Doc, 1.31.95, p.6). And just like Sandra, Jeff emphasized steps for implementing criterion-referenced evaluation in classroom practice through discussing and negotiating criteria with students.

"Kay," the Intermediate Coordinator, chaired a "Teachers-As-Researchers" group, of which Anne was a
member, to explore issues associated with criterion-referenced evaluation. During one meeting, she distributed a handbook to group members entitled, "Criterion-referenced Evaluation - The Road to Success is Always Under Construction" (FN, 2.11.95; Doc, 2.11.95). Within this context, it is useful to recall that Sandra Evans ended an earlier workshop on criterion-referenced assessment with the expression, "The road to success is always under construction" (FN, 11.3.94, p.79).

Like Sandra's and Jeff's handouts, Page 6 of Kay's booklet detailed steps for using criterion-referenced evaluation in the classroom. Steps #4 and #5 suggested that teachers, "involve students, when appropriate, in establishing criteria (and), inform students about the criteria their work will be evaluated against prior to the learning activity" (Doc, 2.10.95, p. 6).

During an M.Ed. course in July 1995, Anne continued her efforts to make sense of criterion-referenced evaluation by reviewing literature associated with the term, criterion-referenced. She examined how the term was used in Robert Glaser's (1963) discussion of criterion-referenced measurement,
the Sullivan Royal Commission (British Columbia Ministry of Education, 1988), and current BC Ministry assessment policy (British Columbia Ministry of Education, 1994a; 1994e; 1994i). She decided that an overarching theme of policy documents and literature was that assessment and evaluation should be directly associated with curriculum and instruction (Doc, 7.25.95).

Anne pointed out that teachers were confused over how to interpret the term, criterion-referenced evaluation.

...I think that people's perceptions regarding what criterion-referenced evaluation means are different. For example, are we talking about explaining criteria to students before they begin an activity but we haven't ensured that they understand exactly what we mean by criteria? Are we also talking about establishing criteria with students? (Inter, 5.22.95 p. 15)

...presently, I wonder if criterion-referenced evaluation must always involve discussion with students, and if criterion-referenced evaluation involves a teacher setting criteria and students simply seeing criteria before beginning an activity. (Inter, p. 17)

Although Anne preferred to discuss and negotiate criteria with students, she said that the process was time-consuming (Inter, 5.27.95). Therefore, she was unable to use the strategies as often as she might have liked. She pointed out that she wanted to make time to
use other procedures and practices, like tests, that she did not associate with criterion-referenced evaluation. (Inter, 5.28.95).

Assessment-Evaluation Procedures and Practices

Anne used multiple approaches to assessment-evaluation in her classroom practice involving gathering, marking, returning, and filing of students' assignments in portfolios, criterion-referenced evaluation activities, teacher observation, homework, and letter grades and structured written comments. She thought that different skills and content areas called for different approaches to collecting information and making judgements about student progress. Anne believed that multiple approaches gave students a variety of opportunities to demonstrate understanding. Within this context, students had greater access to achievement. Anne also wanted students to gain experience with assessment-evaluation practices commonly used by teachers in junior high and high school.

Marking. Anne's classroom assessment-evaluation practices involved daily collection, marking, returning, and sometimes filing of students'
assignments in portfolios. Assignments took the form of reading responses to literature, journals-writing, notes, maps, stories, tests, problem-solving work sheets, student self-evaluations, science experiments, presentations, and art-works.

I think people should use a range of assessment tools...because you should be teaching kids a range of things. (Inter, 5.28.95)

Each day students worked on assignments that they were required to complete within certain time frames. Anne maintained a list of completion dates for assignments on the blackboard at the front of the classroom. Students handed in completed assignments for marking according to the schedule.

In some cases, work was due the same day that it was assigned. Most due dates crossed days and weeks. Students usually had three or four days to complete tasks in math, language arts, social studies, and science. They had several weeks to complete projects. Students stayed after school to finish assignments. Anne notified parents of those students who needed to spend extra time on their work at home.

Anne assigned a mark based upon how effectively a student demonstrated her or his understanding of
performance criteria. Criteria varied according to expected learning outcomes for a subject. When students worked in groups, members of the same group received the same mark for an activity.

Anne carefully considered students' self-evaluations when she marked assignments. When a student assigned work a significantly lower or higher mark than did Anne, she discussed the discrepancy with the student. In some instances, Anne changed her mark after speaking with a student.

Anne interpreted marks as numerical, figurative, and written symbols that represented the degree to which performance criteria were met. Her choice of symbols depended on the nature of the assignment.

You have to change how you're going to mark assignments according to what it is you're doing. (Inter, 5.28.95)

She explained that she used different approaches to marking assignments based upon the "objectivity" and "subjectivity" of the subject matter (Inter, 5.28.05). She interpreted a subject as objective when its content could be tested on the basis of right and wrong answers. Computations skills and spelling were examples of objective subject matter. In contrast,
when a subject did not lend itself to tests involving right and wrong answers, she thought that it was subjective.

Anne described reading as subjective because students approached their assignments from many different angles. She thought that her understanding of criterion-referenced evaluation reduced the subjectivity of language arts assignments and social studies projects because she and the students jointly negotiated performance and evaluation criteria. In this way, Anne did not feel solely responsible for deciding what counted as average, below-average, and above-average achievement. By discussing criteria with students she solved the dilemma of meaning-making in relation to ideas like creative, interesting, complex, and neat.

Anne marked creative writing assignments and projects in language arts and social studies by means of written comments and numbers from 1 to 4. Students also chose a mark and wrote supporting comments for their work. As previously stated, Anne spoke with students about their self-evaluations within the context of a teacher-student conference when there were
large discrepancies between marks.

Anne marked tests in spelling and math out of totals that ranged from 10 to 50 marks or more. She used tests from two teacher's manuals associated with a grade 5 and 6 program, and tests from the handbook of a commercially prepared spelling program.

When reviewing students' plans for social studies research, Anne indicated her approval by initialing the planning sheet. When plans were endorsed in this manner, students were permitted to proceed to the next stage of research.

Anne's approach to marking also involved writing comments on assignments in story-writing, art and social studies. During a lesson regarding the "Voyage of the Mimi," Anne wrote comments concerning how thoroughly students responded to videos and print material. She pointed out where and how individual students might expand their ideas, and noted errors in grammar and spelling. With the exception of their journals, students were expected to submit a second draft of writing that was free of spelling and grammatical errors.

Anne collected and marked students' assignments at
different times. Sometimes she collected information and judged student performance while a student engaged in a problem-solving activity. During research presentations, we jointly determined marks and wrote comments on a student's criteria worksheets. Worksheets indicated performance criteria for each category or stage of an assignment. Students received copies of the sheets when they began an activity. Before giving presentations, students returned their sheets to Anne with self-evaluation sections completed.

Anne also marked students' assignments before and after school, during recess, and at home. Occasionally, students marked their own and/or each other's work, and then submitted it to Anne for review. For example, students completed end-of-unit tests from their math textbooks and checked the accuracy of their responses against answers in the teacher's manual.

Anne stored her marking in several ways. She kept an ongoing record of students' marks in a mark book. She also collected work samples with marks and/or comments and filed them in students' portfolios. Work samples took the form of artwork, projects, work-sheets, and journals. Students sometimes kept
assignments in duotangs which Anne regularly called in for review. It was difficult for Anne to find enough storage room for the work she collected over a term.

Anne's approach to marking involved procedures and practices for gathering information about student progress, making judgements about this information, and storing work samples with marks and/or comments in student portfolios for ongoing consideration. We agreed that I would gain insights into her classroom assessment-evaluation procedures and practices by helping her with marking. In addition, I would better understand how she constructed letter grades for the report card.

Criterion-referenced evaluation. Anne saw criterion-referenced evaluation as strategies for collecting and making judgements about information regarding student progress based upon activities wherein students participated in setting criteria. She liked criterion-referenced evaluation because she enjoyed discussing and negotiating criteria with students. She explained that, "...I think...trying it with the kids, you prove to yourself it's very valuable" (Inter, 5.22.95, p. 23).
By mid-July 1995, Anne's research and reflections on her understanding of criterion-referenced evaluation led her to make the following observations.

During the past school year, my students and I have worked collaboratively to implement criterion-referenced evaluation. The value of this process is that we have developed a shared understanding of curriculum expectations and criteria for assessment and evaluation. (Doc, 7.25.95, pp. 7-8)

She added that discussing and negotiating criteria with students resulted in several outcomes that contributed to shared understanding. For example, achievement levels often improved when students dialogued about the significance and meanings of performance criteria. Anne pointed out that students did not always understand her expectations for an assignment when she told them what the criteria were; even when she wrote them down. Ongoing discussion and negotiation among her and students clarified the meaning of criteria.

Anne: It should be clear to students what you expect of them to get a particular mark. The discussion that you have with students is extremely important. I can write something on the blackboard, overhead, or a work sheet but it may not be clear to students what I expect of them. (Inter, 7.27.95, p. 4). This never seems to happen with criterion-referenced evaluation because we've had that discussion and everybody understands what everybody is talking about.
Anne came to believe that criterion-referenced evaluation involved sharing her power to name and define performance criteria and standards of achievement with students (Doc, 7.25.95). She thought that sharing her power with students in this way contributed to a working environment that promoted mutual respect and shared understanding (Inter, 7.27.95).

Anne commented that some teachers were concerned that students might lose respect for them if they shared their power in this way. However, she said that she had not found this to be the case.

I'm the teacher but I don't know everything. There are things that I haven't thought of that they think of, and I think sharing your power with students really means guiding students. I feel comfortable doing that. Students are probably more cooperative too. They certainly do not lose respect for a teacher with this approach. In fact, they probably respect a teacher more for respecting them in this manner. (pp. 5-6)

I believe that the process of negotiating criteria means that a teacher shares her power with students to set performance and evaluation criteria. Sharing my power in this way contributes to the development of mutual respect in the classroom and creates a supportive working environment. (Doc, 7.25.95, p. 10-11)

Anne suggested that mutual respect among a teacher and her students was an important factor in developing
shared understanding (Inter, 5.22.95). She believed that mutual respect was achieved in an environment where participants felt "safe" to express their opinions and ask questions.

...you need to have a safe environment in your classroom where students can feel comfortable to share ideas and ask questions. Children need to feel that a teacher respects their ideas and opinions to feel safe. Moreover, they do not feel safe if other students criticize them or laugh at them. A teacher needs to prevent this from happening. (Inter, pp. 18-19) Children have wonderful minds and we need to appreciate and respect that fact. (Inter, 5.28.95, p. 10)

Anne added that students' knowledge and interests were recognized through collaboratively developed curriculum (Doc, 7.25.95). In this way, she empowered students' knowledge by recognizing it as an important source of criticism.

Anne outlined her approach to negotiating criteria with students in her reflections on a Language Arts lesson (Inter, 7.25.95). The objectives of the lesson were to reinforce writing skills in terms of using paragraphs and writing from another viewpoint. Anne introduced the lesson with a video. After the students watched the video, Anne instructed the students to write a story from the point of view of one of the characters. Before the students started writing, they
discussed and negotiated the performance or evaluation criteria for their stories with Anne.

Anne: After we watched the video, the students and I negotiated criteria that would be used to evaluate their stories. I initially proposed some criteria that I felt were important. Then I asked students if they felt that my criteria were important, and if revisions were needed. At the end of this discussion, students decided which criteria should be used to evaluate their stories. Then, I used guiding questions to help students determine three or four more criteria for the activity. While they discussed criteria, students shared opinions regarding each other's points of view. We made a group decision regarding which criteria should be used to evaluate the assignment. (Inter, 7.25.95, p. 8)

Each student received a work sheet that listed the collaboratively developed evaluation criteria for the assignment (FN 3.31.95; Doc, 4.10.95), (Appendix B). The criteria were: "must follow story line of video; must imagine the gaps in the story and write with creativity and description; must write from one character's point-of-view and show the character's personality; must have a beginning, middle, and end; must use new paragraphs for new ideas; must use correct mechanics of writing."

Anne and the students marked the stories out of 4, and wrote a comment to explain their choice of mark. A story that met all of the criteria received a mark of
"4". A score of "4" represented the standard, "powerful or highly effective." A story that met the majority of criteria received a score of "3" which meant "capable, confident." A score of "2" meant that the story demonstrated "partial, developing" understanding of criteria. "1" meant that the story indicated a "limited, undeveloped" understanding of criteria. Criteria for "powerful," "capable," "developing," and "undeveloped" were on charts on one wall of the classroom (Doc, CH-94), (Appendix C).

Within another context, Anne and her students discussed and negotiated criteria for a Dr. Seuss-style booklet. Anne pointed out that the dialogue that occurred among the students and herself ensured that "...everyone...knew what everyone else was talking about...." (Inter, 5.28.95). Following is an excerpt of the discussion (FO, 5.26.95). Some of the students are identified by a pseudonym. Others are represented by "S."

Anne initiated discussion by asking students what criteria should guide the development of their books. When a student responded that "rhyme" ought to be a criteria, Anne encouraged the group to examine what
rhyme meant in relation to Dr. Seuss' work.

Anne: What do you think should be the criteria for marking our Dr. Seuss-style books?
S1: It should be like Dr. Seuss' style.
Anne: Okay in (writes on chart paper "In Dr. Seuss' style").
Anne: Okay...tell me what it means by in Dr. Seuss' style.
S2: Rhyming.
Anne: Okay rhyming (writes on chart). Now maybe we should explain exactly what we mean by rhyming. Which parts rhyme? Is it just the end? What do you mean by the term rhyming? (FO, 5.26.95)

Anne's questions concerning "rhyme" triggered and guided students' explorations of criteria. David responded to Anne's questions by suggesting that Dr. Seuss used rhyme according to the style of a particular book. She proceeded to seek consensus from the group by asking students if they understood the criteria of rhyme.

Anne: Okay so does everybody understand what we mean by rhyming? (Students nod heads in agreement) Okay that's Dr. Seuss style and we said that's one of the criteria you're being marked on (is) that you are able to write in Dr. Seuss' style. Therefore you must write so that the last words rhyme with each other or the words in the sentences rhyme. Now David how about if you get an example of each of those kinds of books to read to them right now. (David reads "Hop on Pop" to class) (FO, 5.26.95)

The discussion continued with students sharing opportunities to negotiate criteria according to their experiences with Dr. Seuss materials.
Anne: ...What else do you think should be a criteria for our Dr. Seuss-style book?
Meghan: Put a picture with it?
Anne: Okay, put a picture with it. Okay I think we need to work on the picture part....I guess you should think about Dr. Seuss' art. I think you need to describe his art. What kind of artwork does he have, I mean are they fancy elaborate pictures, or are they quite plain pictures, are they dull, are they colourful? I think we need to zero in on exactly what kinds of pictures he does.....
S3: His is not real.
S4: Yeh, his is imaginative and make-believe, but that's more realistic and
S5: Real-life.
S4: Yeh, and they look like real people.
Andrew: Dr. Seuss art is like it's shaggy. All his people are like all slouching and his trees are all bent over...they look funny.
S6: He draws what he writes like the "Cat in the Hat".
Anne: Okay and Casey would like to say something.
Casey: His pictures are very simple so they're very unique how he makes the pictures but the colours are very simple.
Anne: And Jeremy?
Jeremy: There's main characters in the drawings for like the whole book.....(FO, 5.26.95)

Anne prompted further development and negotiation of criteria by piggy-backing on a student's reference to the backgrounds of Dr. Seuss' pictures. This time she sought consensus by having the group raise their hands if they agreed with criteria.

Anne: Somebody said...that the background is always very simple. Would you like that to be a criteria? Simple background. How many people would like that for a criteria (Several hands go up). Okay how many people don't think that should be a criteria?
Jeremy: Can you have a detailed background?
Anne: Okay what do you think....? How many people think no, it shouldn't be detailed. Let's take Angela's comment.
Angela: Okay, where's "Hop on Pop"? Look at this it's just like...there's a tent and him and little spikes....and then it's just plain white.
S8: There's no background it's just the picture.
Anne: Paul had one comment.
Paul: He just put what he needed just like a tent or whatever.
Anne: Should we say then that the background is simple?....Hands up if you agree with that. Okay that looks like the majority of, most, just about everybody, more than 3/4's. (FO, 5.26.95)

Although Anne pointed out that the majority of students agreed with criteria, she did not impose meaning by resorting to majority rule. Instead, she asked for a show of hands to ascertain the group's degree of agreement and disagreement. Anne asked each student who disagreed, or failed to raised her or his hand, to explain their concerns. She used these concerns to readdress the criteria under consideration.

Anne concluded that negotiating criteria with students led to increased interest and enthusiasm in their work (Inter, 5.28.95). She thought that students demonstrated a heightened sense of ownership and responsibility towards their assignments.

Teacher observation. Anne systematically observed how effectively students worked on tasks individually,
with peers, and during teacher-directed lessons. Sometimes she recorded her observations. Other times she simply made mental notes. Anne expected students to remain focused on a task until it was completed. During group activities, students were encouraged to work collaboratively. For a teacher-directed lesson, they were required to pay attention and participate when appropriate.

When Anne saw students exhibiting off-task behaviour, she immediately reminded them of their responsibilities. For example, students were permitted to talk, but talk was to be task-oriented and conducted in soft tones. Students could only leave their desks to walk around the room when they needed to use the stapler or three-hole punch, to acquire or deposit scrap paper from the recycling box, to sharpen pencils, and when they needed assistance from Anne. Anne quickly saw if students were engaged in inappropriate behaviour and spoke to them accordingly. When she was uncertain why a student was engaged in a particular activity, she asked the individual to explain what they were doing before she judged their behaviour.

In addition, Anne made mental notes of student
performance and behaviour while they worked. Although she did not have time to record all of her observations, they contributed to her intuitive understanding of how a child was doing.

Homework. Anne reviewed student progress by checking homework books at the end of each day. Homework consisted of assignments that students were required to finish in subjects like language arts, mathematics, science, social studies, and art. A student was not dismissed at the end of the day until her or his homework book was signed by Anne. Anne encouraged parents to sign the homework book to indicate their awareness of work completed, and work due.

When Anne saw a student lagging behind schedule, they stayed after school to work with her until 3:30. Students also worked on homework during class time after they finished an assignment from a particular lesson. Some students came to school at 8:30 a.m. to finish work before classes started at 9:00 a.m. Students who finished assignments before due dates did not have homework.

Because homework was not always done at home, Anne
was able to observe students while they worked and provide assistance when necessary. Homework stemmed from unfinished daily assignments, short-term preparation for math and spelling tests, and long-term preparation for projects in math, science, social studies, and language arts. Homework grew directly out of daily activities.

Letter grades and structured written comments. I assumed that by understanding Anne's approach to marking, I would understand how she constructed letter grades and structured written comments for report cards. By the end of the second term, I had not figured out how Anne constructed letter grades and comments for report cards. At the end of the third term, Anne permitted me to observe her write a report card. By that time I was better able to understand her rationale for constructing letter grades and structured written comments because I was familiar with her classroom assessment-evaluation practices.

Watching and listening to Anne while she constructed letter grades and structured written comments was a great privilege. During eighteen years of teaching, I came to believe that letter grades were
at the heart of a teacher's accountability. As a classroom teacher, I often worried that my approach to creating letter grades would fall short of the expectations of administration and parents. In addition, I worried about defending my approach to parents and principals. Consequently, I told no one how I came up with a letter grade. Luckily, no one asked. Writing report cards was a closeted activity when I was teaching. I suspect that this is still the case for some teachers, if not most.

Anne invited me to observe her write a report card one day after school (FO, 6.23.95). She sat at a table covered with portfolios, piles of tests, duotangs, folders, exercise books, projects, and self-evaluations (FO, 6.23.95). In front of her was a blank report card. At her side was a mark book.

We agreed that Anne would not refer to the student by name, and that I would sit several feet away from the table. My location prevented me from seeing a student's name when Anne made notes for her or his report card, and when she referred to her or his work samples. During my classroom career, I learned that a student's report card was off-limits to anyone who was
not directly associated with a child's program. That is, people with the right to know of a student's letter grades were the student, her or his parents, school administration, and those teachers or professionals who were directly associated with the child's program. Although I was directly involved with Anne's students, I did not believe that I had the right to know of students' grades. Furthermore, I did not think that I needed this information. In other words, I wanted to understand Anne's procedures for creating letter grades. I did not need to know who got them.

Throughout the observation period, Anne explained to me what she did and why. At first I tried to write everything she said, but eventually asked if I could use the tape recorder. She agreed.

Anne explained that before she started writing a report card, she reviewed a student's previous reports. She wanted to ensure that she did not repeat general comments she had used in past reports. She also wanted to see where a student's letter grades stood in relation to current performance. When a letter grade went down significantly from a previous report, she addressed the change through detailed structured
written comments. Included in her comments were references to documented evidence to support the change. Anne recorded her observations from the student's previous reports on a sheet of paper. At the top of the paper were headings for each subject area (FO, 6.23.95).

After examining past report cards, Anne reviewed information she gathered regarding the student's progress. This information included samples of work from several piles on the table, and marks in the mark book. She explained that even though each report card took approximately three hours to complete, she wanted to refresh her memory concerning why the student had received certain marks.

It is extremely important to refer to work samples when preparing structured written comments and letter grades for a student's report card. If a teacher relies solely on her mark book she might not recall why a student received a particular written comment or mark. This approach is time-consuming but I think that it leads to more valid judgements regarding a student's progress. (FO, 6.23.95)

When Anne finished reviewing her assessment information, she wrote a general comment concerning the child's progress, strengths, and things that they had done well during the term. Then she addressed their
progress in specific subject areas with an emphasis on strengths.

Anne constructed the student's letter grade and comments for language arts from information associated with five activities. The activities involved reading a biography, researching and presenting a biographical report, writing and illustrating a Dr. Seuss-style booklet, researching and writing a report on an area of interest, writing from the point-of-view of characters in a video story and a fairy tale, and preparing and conducting an interview. Anne explained that all of the activities consisted of skills associated with reading, writing, listening, and speaking.

She added that language arts was completely criterion-referenced evaluation. By criterion-referenced evaluation, Anne meant that she negotiated criteria with students before they engaged in an activity.

All activities, with the exception of the interview, were marked out of 4 with supporting comments. I guessed that Anne totalled the marks and made some judgement about whether the total represented an A or B etc. Wrong again.
I'm thinking about each assignment in terms of, is it a four or a three? I have not recorded if a student received four or three in my mark book because I can readily look at her or his work. Students know that a four is an A and that a three is a B. I do not have this recorded in my mark book. This information is recorded on criterion-referenced forms. In this way, I have all of the criteria for assignments in front of me. If I only looked at my mark book, I would not be able to see the criteria or skills. (F0, 6.23.95)

Anne constructed the student's letter grade and structured written comments from marks and comments on the students' work sheets. Structured written comments included a description of the student's strengths, areas needing improvement, and goals for improvement.

While Anne wrote the letter grade and comments on the report card, she explained that she kept draft copies of the student's work attached to final copies. In this way, she saw what the student needed to improve to arrive at the final copy. Completed work was sent home with draft copies attached so that parents saw the transition from draft to final copy. In this way, Anne assumed that parents had a better understanding of the letter grades their child received.

...if I file only good copies of students' assignments, I might not remember what skills needed reinforcement and areas that I wanted to discuss on the report card. Furthermore, when parents compare a draft copy of their child's work to a good copy of the same assignment, they can
see what skills needed improvement, and the degree to which improvement was achieved. For example, if parents read a teacher's comments on a report card like, 'needs more creativity and description, and be careful to use capitals and periods', and they have only seen corrected copies of their child's work, they wonder what she's talking about. (FO, 6.23.95)

With the language arts section of the student's report card completed, Anne turned her attention to mathematics. While she worked, Anne explained that she constructed the student's letter grade and comments from unit tests and problem-solving activities. Students completed unit tests on computational skills from a commercially prepared math program. Anne acquired ideas for problem-solving activities from a local math professor, math workshops, and commercially prepared problem-solving booklets. She integrated math and science skills in problem-solving activities. Activities involved problem-solving with pattern blocks, geoboards, graphing, and measurement.

The student's marks for unit tests were recorded in the mark book. Anne totalled the marks, converted the score to a percent, and translated the percent to a letter grade based upon standards for letter grades used by high schools. She said that she intuitively made judgements about problem-solving activities
according to a numerical standard of 1 to 4 wherein 1 represented an A, 2 was a B and so on. When I used this strategy as a teacher, I called it eyeballing.

...when you read what they've done you have a very good idea if this is a four, this is beyond a four, if this is a three or less then they really didn't understand that much. Even though I don't mark every single thing on a sheet it is still in my your head. (FO, 6.23.95)

Before she selected a letter grade, Anne reviewed the student's tests and problem-solving work sheets. She added comments concerning the student's strengths, areas needing improvement, and goals for improvement.

Anne determined the letter grade and structured written comments for science from marks and comments on the student's notes, problem-solving work sheets, and tests. Because the program was inquiry-based, Anne said that she preferred not to use paper-and-pencil tests to assess student achievement. Instead, students engaged in hands-on problem-solving activities. They used guide sheets with performance criteria to direct their learning.

Anne told students that they would be marked out of 5 for all of their assignments. She also explained performance criteria for the numbers 1 through 5. Anne told me that science, like language arts, was
criterion-referenced because performance criteria were explicit. She thought that if criteria for an activity were not made clear to students, then the activity was not criterion-referenced.

To arrive at a letter grade for science, Anne totalled the student's marks and converted the raw score to a percentage. She wrote comments regarding the student's strengths, areas needing improvement, and goals for improvement.

Moving on to social studies, Anne explained that she used marks and comments from projects and daily assignments to construct letter grades and structured written comments. Marks for review assignments varied according to the detail and complexity of response that was required. Marks ranged from 1 to 6 and were accompanied by Anne's comments.

Projects were marked by Anne and students according to the degree to which performance criteria were met. Plans for projects were marked out of 8. The written report was marked out of 4. Each of the four information products was marked out of 4. Comments accompanied each set of marks. Marks for projects were totalled and transformed into a
percentage or letter grade that Anne recorded in her mark book.

Anne finished writing her comments for social studies and turned to the section of the report card reserved for art. She explained that the letter grade and comments for art were based upon marks and comments for activities the student completed in language arts, science, social studies, and art class. She added that because she did not have criteria for art, letter grades were "subjective" (FO, 6.23.95).

Anne was uncomfortable giving a letter grade for art for two reasons. First, she was unsure of performance criteria and related performance standards for art because she was not an art teacher. Second, she said that it was not fair to mark a student's art until he or she could specialize in high school.

I think that it is silly to give a letter grade for art until a student goes to high school where they might specialize in the subject. I am concerned for a student who is doing her or his best, but can not draw very well. I think that it's too bad that he or she might receive a C+ or less. I don't give anybody lower than a C+ in art. I do not believe that a letter grade for art is important in elementary school. However, I can understand that a letter grade would be important in high school. Basically, if students are trying very hard and are doing the best they can, what more can I expect? (FO, 6.23.95)
Anne's discussion of letter grades and art is a vivid example of how her interpretation of letter grades was situated in her personal history. She was concerned that her judgement of a student's art would be unfair due to her lack of training in the subject. Consequently, she recommended that students receive letter grades for art when they were taught by art specialists in high school.

Anne was also in conflict over how to mark students who were doing their best but were unable to draw very well. However, she was not in conflict over marking students who tried to do their best in math but were unable to solve problems very well. I recall thinking like Anne when I had to give letter grades for art in grades 5 and 6.

Many people associate art with creativity. I have heard many teachers say that they were not creative, and were therefore uncomfortable judging classroom art. I failed art in high school so I never felt competent to judge student art. Anne too was convinced that she couldn't draw.

I remember her anxiety about drawing during silk painting classes. Anne wanted to do a wall-hanging of
irises. She was going to copy a picture of the flowers from a book. When the instructor told her to go outside and sketch flowers free-hand, Anne pleaded, "but I can't draw!" After several reassurances from the instructor, myself, and two friends, Anne made some marks on her piece of paper. At some point, the instructor took Anne's eraser away because she rubbed out more than she drew. By the end of the course Anne made a lovely wall-hanging of irises. It hangs in her master bedroom. So, when I heard Anne speak about how unfair it was to grade students who tried their best but couldn't draw, her words echoed from the silk-painting class.

By now both Anne and I were getting tired. It was nearly 6:00 p.m.. Anne was almost finished writing the student's report card. All that remained for her to do was add grades for french, physical education, and music from class lists that she had received from the respective subject teachers. Anne walked over to the filing cabinet and grabbed a paper bag from the top drawer. Then she returned to the table and sat down. Anne dug a sandwich and banana out of the bag and began to eat her supper. While eating, Anne explained that
at the end of each term students organized samples of their work in portfolios in preparation for student-centered conferences. Work samples were selected by her and students. She remarked that she liked student-centred conferences because they gave students opportunities to explain to parents how daily work was linked to letter grades and comments on their report cards. In addition, discussing portfolios with parents led children to take greater responsibility for their learning. Before leaving, I thanked Anne for the privilege of learning how she created letter grades and structured written comments. I did not ask how late she planned to work. I already had a pretty good idea.

Collegiality

Anne explained that having opportunities to share information with colleagues was important to her classroom practice.

The teacher-as-researcher group provided me with opportunities to explore and share ideas with colleagues. (Inter, 7.27.95, pp. 13-14)

She pointed out that she was not comfortable talking about her classroom practice with colleagues at Westheights. She worried that they might think that she was "bragging."
I do not want to go into the staff room and tell my colleagues that I am doing all of these wonderful things with students. They might not be interested in these things. Furthermore, I do not want my peers to think that I am showing off or bragging about myself. I am doing criterion-referenced evaluation because I am really interested in it. (Inter, 5.27.95, p. 14)

Anne expressed the need for shared understanding among teachers and administration regarding the policy change. She pointed out that although the Ministry stipulated uses of terms and their meanings in policy documents, people interpreted meaning in different ways according to their experiences.

Until you have talked to somebody, you do not really know what they are talking about or what they understand. (Inter, 5.22.95, p. 29)

Interviews with several of Anne's colleagues supported her comment that people understand terminology differently (Inter, 5.29.95; Inter, 6.11.95; Inter, 6.26.95; Inter, 6.27.95). When I asked "Vera," the Special Education Resource teacher, about criterion-referenced evaluation she said that she favoured curriculum-based assessment practices. These practices involved teacher observation, working with students, and ongoing judgements concerning how effectively students demonstrated their understanding of curriculum criteria.
I don't believe progress is made...in putting a child in 'oh you read at a grade 4 level. Let's find some grade 4 material....to me it's reasonably irrelevant at what grade level they read. Whenever possible I like to work with the curriculum...It's very empowering for a kid to do okay with what everyone else is doing. (Inter, 5.29.95)

At no time during our interview, did Vera use the term, criterion-referenced evaluation.

"Pam" and "Jan," the Learning Assistance teachers, said that they were not sure what criterion-referenced evaluation meant (Inter, 6.27.95). They added that they used curriculum-based assessment practices and standardized tests to determine students' needs.

Pam: ...we're a little foggy about it at the moment so we don't know how it's going to affect us....
Jan: ...we had one, was it a quarter of a day, or half of a day, that was on criterion-referenced assessment and it felt extremely rushed and I think people came away from it saying, 'I already knew that, what's that all about?' (Inter, 6.27.95)

The workshop to which Jan referred was facilitated by Jeff, the District principal. It is useful to recall that although Jan referred to criterion-referenced assessment, Jeff talked about criterion-referenced evaluation (FN, 1,31, 95). Neither Pam nor Jan used the term, criterion-referenced evaluation, during our interview.
"Clint," the Physical Education teacher, talked about the importance of students being aware of evaluation criteria throughout all of their activities (Inter, 6.11.95). He favoured performance evaluation wherein students actively demonstrated their mastery of skills. He referred to the BC Premier Sports system for standards of achievement in various sports. Clint associated criterion-referencing with "types" of evaluation that included teacher, student, and peer evaluation. He did not use the term, criterion-referenced evaluation.

Well I think as I said at the start (of the interview) was my focus on learning... anything that we do, that the child knows that there is an evaluation process and that you don't spring it on them at the end. That you let them know at the beginning what you're going to be testing and how you're going to go about testing. Whether it's going to be student evaluated or teacher evaluated and that it is ongoing so the child doesn't say in six weeks he's going to be looking for something. That he is looking for things all the way along....so criterion-reference I think is an important concept but I also think that, that the other types of evaluation of that are in there....(Inter, 6.11.95)

"Doris," the music teacher, also did not use the term, criterion-referenced evaluation. She explained that although the phrase criterion-referencing was "new" to her, she interpreted it as a useful strategy
for strengthening teacher accountability.

Specifically, when parents wanted to know why their child received a particular letter grade, the teacher was able to respond by showing parents criteria for particular standards of achievement.

...I've been out of teaching for a little while....The phrase (criterion-referencing) was new to me....But that helps a lot in making the teacher accountable...if you lay out your criteria and say this is the criteria and then you can just match it with the achievement, and it helps you get around those widely held expectations a little bit, because it just makes it very specific. So if a parent says 'where did this B come from?', you can say 'well this was what they had to do.' They had to do this assignment...and maybe there were no tests this time, but these were the criteria I used, and this is how the kid did. (Inter, 6.19.95)

Doris' discussion implies that parents might interpret tests as more explicit representations of criteria for making judgements about student progress than practices associated with criterion-referencing.

Whereas Anne understood criterion-referenced evaluation as, in part, an instructional strategy to enhance learning, Vera neither used the term nor any of its derivatives. Pam and Jan used the term, criterion-referenced assessment, but could not provide a definition. Clint used the term, "criterion-reference," as a type of evaluation associated with
student and teacher evaluation. Doris saw "criterion-referencing" as a strategy to enhance accountability with parents. From this experience, I learned that teachers could not agree upon what the term, criterion-referenced evaluation, should mean nor how it should be used. In addition, there was no consensus regarding uses of the terms, criterion-referencing, assessment, and evaluation.

Working Conditions

Anne pointed out that she did not have enough time to accomplish her teaching goals and pursue professional reading, research, and professional development (Inter, 5.28.95; 7.8.97). She explained that the workload represented by marking students' assignments and creating letter grades and structured written comments for report cards was often overwhelming.

Anne regularly worked many hours of overtime to accomplish these tasks; time that she wanted to use on other important professional activities, and with her husband and two sons. Continuous overtime led to fatigue. Anne described a typical day of teaching in the following manner.
I come to school at about 8:00 in the morning. Probably it's my only really relaxing time of the day. I interact a lot with the kids and talk to them as they come in, but then of course you start teaching. I only usually take about a half an hour for lunch time and use the rest for working. And after school I am usually here working until about 6:00, that's the average time and it's spent on preparing and marking. And then I go home and I'm still doing marking or preparing, or thinking about new things...planning new things...going to workshops and attending meetings and things like that and I hardly ever watch T.V.. Usually I'm working until I go to bed at night....even though I put in that much time...I still am behind in my marking....you just can't keep up with the marking....It's very frustrating. (Inter, 6.23.95, pp. 22-23)

She added that she tried to spend time each day with her husband and two sons.

Anne explained that she neither had sufficient time nor opportunities to collaborate with colleagues regarding education policy and classroom practices. She felt isolated in her efforts to make sense of the policy change in her classroom practice. Anne pointed out that she lived in a room filled with children five days a week. She explained that even though she loved kids, she still needed professional time with colleagues. Because these opportunities were not available, Anne encountered difficulties with making sense of the policy change; difficulties that contributed to feelings of frustration and anxiety.
Conclusion

I interpreted Anne's classroom assessment-evaluation practices as mediational means for supporting learning and developing metacognitive skills. That is, these practices served as means for monitoring student progress, increasing student achievement, and teaching students how to take control of their learning.

Anne saw criterion-referenced evaluation as, in part, an instructional strategy that involved student participation in setting criteria. She came to believe that learning and achievement increased when she collaborated with students in setting performance criteria. That is, Anne thought that sharing her power or authority to name and define performance criteria with students was central to their learning and achievement. Within this context, students had increased control over their learning, and the means by which marks and grades were generated.

Anne's understanding of criterion-referenced evaluation as a classroom practice wherein learning and achievement increased when students were involved in setting performance criteria recalls Vygotsky's (1978;
"zone of proximal development." Vygotsky (1978; 1986) saw a child's learning in the zo-ped as mediated by assistance from a more capable individual. Within this context, the purpose of mediation was to "meet" adult expectations regarding performance.

Consistent with Vygotsky's emphasis on learning mediated through assistance, Anne helped students learn to identify and negotiate meaning for performance criteria. Unlike the relationship between a child and more capable individual in the zo-ped, Anne shared her power with students to create meaning for performance criteria and achievement standards. That is, she did not expect students to meet her criteria for learning and achievement without their input; albeit, within the parameters of Anne's overarching goals for an activity. Anne came to believe that more equitable power relations positively mediated learning and increased achievement.

Although Ministry policy regarding criterion-referenced evaluation recommended that teachers involve student participation in setting criteria when appropriate, Anne decided that criterion-referenced evaluation was a classroom practice wherein she always
shared her power to set criteria with students. Anne's decision to modify policy definitions of criterion-referenced evaluation recalls Wertsch's (1991; 1995) observation that mediational means are frequently changed to increase their efficiency, and issues associated with power often guide peoples' interpretations and uses of mediational means.

Although Anne thought that criterion-referenced evaluation improved learning and achievement, her understanding of the idea generated difficulties for her classroom practice. She pointed out that involving student participation in setting criteria was time-consuming. She needed more time to develop a rationale to determine when it was appropriate to involve students.

Anne's understanding of criterion-referenced evaluation problematized her understanding of criterion-referenced letter grades. She reasoned that because criterion-referenced evaluation involved negotiating criteria with students, criterion-referenced letter grades were constructed from information generated by this practice. Because Anne did not have enough time to use her interpretation of
criterion-referenced evaluation in all subjects, she worried that her letter grades were not entirely criterion-referenced.

Anne's struggle to understand criterion-referenced evaluation and criterion-referenced letter grades reflected her need for control over her practice. Her accountability to herself, students, colleagues, parents, and administration, was contingent upon her authority and capability to adopt, modify, generate, integrate, and even reject policy discussions regarding assessment-evaluation in her classroom practice. Anne's power and capability to make sense of the policy change were constrained by working conditions wherein she had limited time and opportunities to collaborate with colleagues.

I interpreted the intersection of Anne's personal history with her understanding of the policy change in her classroom practice, from the standpoint of Bourdieu's (1977) concepts of habitus, field, and practice. Within this context, Anne's tacit understanding of assessment-evaluation represented her "feel for the game" (Bourdieu, 1990).

Anne's professional judgement and classroom
practice reflected practice that, as Bourdieu observes, does not draw attention to itself and is even "...unaware of the principles that govern it and the possibilities they contain; it can only discover them by enacting them, unfolding them in time (p. 92)."

Hargreaves (1990) points out that the needs and responsibilities of teachers generated by the dynamic culture of classroom life mean that policy might not be relevant to practice.

...teacher needs and demands generated from the particularities of the context may obstruct, undermine, or redefine the purposes built into new administrative procedures and the time designations and allocations which accompany them. In this strained juxtaposition of monochronic and polychronic time-frames can be seen much of the reason for the apparent failure of administratively imposed reforms in education. (Hargreaves, 1990, p. 311)

Like Hargreaves (1990), Bourdieu (1990) contends that "scientific" or second-order theories of practice obviate the role of real time in constructions of practice.

...because it is entirely immersed in the current of time, practice is inseparable from temporality.... It's temporal structure...is constitutive of its meaning.... Science has a time which is not that of practice. For the analyst, time disappears: not only...because the analyst cannot have any
uncertainty as to what can happen, but also because (s)he has the time to totalize, that is, to overcome the effects of time. Scientific practice is so detemporalized that it tends to exclude even the idea of what it excludes. Because science is only possible in a relation to time which is the opposite of that of practice, it tends to ignore time and so to detemporalize practice. (p. 81)

Only by working collaboratively with Anne, was I able to gain some insights into the elusive, transparent, and ubiquitous qualities of her classroom practice. In the next chapter, I describe and interpret how Anne's students understood and responded to classroom assessment-evaluation practices.
CHAPTER 6

"the big picture isn't always the important picture" (David, Inter, 6.2.95).

THE CHILDREN

Children And The Policy Change

Children were the objects of educational policy. They were the ones being marked and graded, and they were very clear about this. Anne's students were the most sensitive and severe critics of the policy change.

Marks and grades were key stakes in the lives of students. Most students interpreted tests, daily assignments, self-evaluations, portfolios, and projects as means by which marks and grades were generated. They exchanged marks and grades for identity, peer group acceptance, praise, clothing, money, food, compact discs, school supplies, and promotion to the next grade. Students saw themselves not only in competition with peers for marks, grades, and identity, but also in competition with teachers' and parents' expectations regarding their achievement.

I begin this chapter by exploring how children constructed meanings for marks and letter grades.
Following this discussion, I examine their understanding of assessment-evaluation practices in terms of tests, self-evaluations, criterion-referenced evaluation activities, and student-centred conferences and portfolios.

Marks and Letter Grades

Most students wanted high marks and letter grades because they had high exchange value on competitive school and home markets. In contrast, they thought that structured written comments had low exchange value.

Letters. Some students liked the series of letters, A-B-C-D, that Anne used for grades better than the series of letters, O- (outstanding), G- (good), S- (satisfactory), U (unsatisfactory), that teachers used in their previous year's class.

Oscar: ...I like the symbols of the letters better....and it also kind of sounds better than like 'I got a B in French instead of an S+....I think like a A and B and stuff, that was used a long time ago, before, and then like S+ and stuff was used, started to be used a little while ago. (Inter, 5.9.95)

For Oscar (Gr.6), the series, O-G-S-U, did not qualify as letters within the context of letter grades. He explained that the letters A through D had a longer
history of representing student achievement and therefore more credibility.

Just like Oscar, Jeremy (Gr. 5) preferred the series of letters A, B and C. He explained that because they were at the beginning of the alphabet, their hierarchical significance increased their exchange value.

Jeremy (Gr. 5): ...last year...we got O, O for outstanding, G for good, and a few other things like that....I like A, B and C better....when it has the O and G it just doesn't seem like a report card....if you got an O you wouldn't feel as good, if you got an O rather than getting an A....Cause A is higher in the alphabet....(Inter, 5.8.95)

Greg also used a hierarchical model to interpret letter grades.

R: What do the letter grades tell you?
Greg (Gr. 5): ...how good I did my work....because they (letter grades) go in a certain order ....starting at the beginning of the alphabet it goes down to the last letter of the alphabet. (Inter, 5.16.95)

"Summer," a Gr. 6 student, was on a modified program. She said that she felt "bad" when she did not receive letter grades on her report card for some subjects (Inter, 5.5.95). Summer interpreted the absence of letter grades on her report card as a sign of a lack of achievement, and became for her a negative judgement of her as a person.
R: How did you feel about your last report card?
Summer: Bad....Because I did really bad.
R: So is this report card different from the last one?
Summer: Yeh, it's different.
R: How is it different?
Summer: Well, like the first one....I got like all really really bad marks. Some of them I didn't even get any marks because I was so behind. And then this...(time) I got lots of marks, only about three that I didn't get any marks on....
R: How do you feel about not getting any marks?
Summer: I felt really bad so I kind of started trying really hard to get marks.

The absence of letter grades on Summer's report card did not protect her from feelings of failure. The omission made her feel worse since she interpreted the absence to mean that her performance did not even merit an F.

Students assumed that letters A through D represented a hierarchical range of performance standards. They also looked at the back of the report card for definitions of letter grades (Doc, RC-95), (Appendix D). Some students used classroom wall charts to interpret letter grades in terms of categories for powerful, capable, developing, and undeveloped work (Inter, 5.9.95; Doc, CH-95) (Appendix C). Children also adopted their parents' interpretations of letter grades.

Oscar used report card definitions and
classroom wall charts to interpret letter grades. Just like Greg and Jeremy, he saw letter grades as a hierarchy of achievement.

Oscar (Gr. 6): A means excellent, 100% effort, perfect work I guess, um B is capable, confident, very good, like pretty good work, um C is average, satisfactory work, like most of it completed but not all of it but most, and just a C is a little bit lower, like just plain satisfactory, and C— is unsatisfactory, and then D is really low, like really bad. (Inter, 5.9.95)

...the back of the report card tells you what the letters mean.

Cassidy (Gr. 5) based her understanding of letter grades on her parents' interpretations. She explained that her parents saw letter grades as a form of rating.

It's sort of like rating....a standard ....like 86 to 100 is an A and then 85 to like 78 or something is a B....

R: How did you figure that out?

Cassidy: My mom and dad seem to act like it's a rating (to) see if I'm doing well enough. (Inter, 6.23.95)

Exchange value. Children exchanged marks and grades for positive relationships with teachers and parents, membership in the peer group, increased self-esteem, career aspirations, clothing, food, money, school supplies, and for opportunities to stop doing activities that they did not like. In the following excerpt, Brett, Matt, Summer, and Buffy describe the exchange value of marks and grades in terms of
educational and career opportunities.

Brett (Gr.6): I just know I need an education to do the things that I would like to do....I'd like to be maybe a marine biologist. (Inter, 5.16.95)

Matt (Gr.5): The report cards help...to get (you) into a good university.... ...that (education) will help you get a job, just like being an accountant or a lawyer. (Inter, 4.27.95)

Summer (Gr.6): ...when you want to go to like a good school or a university or something you have to have good marks to get into it sometimes.... ...if you get good marks you get a good job when you grow up....(Inter, 5.5.95)

Buffy (Gr.5): ....it (report card) tells me the marks that I got and if I could try harder and so it tells me if I'm not doing good in school or if I am....later on in life....I might want to go into medicine or like I always wanted to be an actress and stuff so I might want to go into acting and drama or something, or...sports....(Inter, 6.5.95)

These children thought that good marks guaranteed access to good universities, and that university education led to professional employment.

Students were concerned with the power of letter grades to affect relationships. Jeremy (Gr.5) explained that letter grades influenced the degree of respect shown to children by teachers, parents, and peers.

Jeremy: ...if you get good marks then people take you more seriously and encourage you and they treat you well, with respect and stuff....If you got a D everybody's going to say you could have
done way better than that, things like that. If you got an A (they) say 'congratulations.'
R: ....like who would treat you better?
Jeremy: Teachers sometimes.
R: Anybody else?
Jeremy: My friends and parents. (Inter, 5.8.95)

David commented that parents might label a child as capable or incapable based upon letter grades on a report card. Moreover, parents might become very upset with a child if he or she received low grades. David concluded that letter grades were potentially damaging to a child's self-esteem and relationships with parents.

David (Gr.6): ...if you don't get an A for your ability then your parents are going to think that you don't have an ability. They'll think you're doing terrible cause this is what you've gotten...and they're very upset with you and that makes you feel terrible cause...you've worked so hard for what you've done. (Inter, 6.2.95)

Greg (Gr.5) pointed out that when he did well on his report card he was praised by his parents (Inter, 5.16.95). Meghan, Martin, and Matt recalled that parents rewarded high letter grades with food, clothing, CD's, and school supplies.

Meghan (Gr.6): if I get a (good) report card they'll take me out to dinner, or this term my mom bought me a shirt.... (Inter, 4.28.95)
Martin (Gr.5): (is rewarded)...only if I get two or three A's.....Then I get some useless stuff like pencils and erasers. (Inter, 5.16.95)
Matt (Gr.6): I know some people, they get a CD
for every A they get... Well, what I get is every time I get an A I get to choose a dinner, what kind of dinner we're going to have, like that my mom can make. Like barbecued chicken, I really like that... teriyaki chicken...(Inter, 4.27.95)

Rick (Gr.6) said that his parents told him he would be allowed to quit "Kumon," a privately funded math class, if he received straight A's on his report card with the exception of Physical Education (Inter, 6.2.95). He recalled receiving $50.00 for a report card of straight A's when he was in grade 5. Rick said that he liked receiving rewards for letter grades depending on what he could get (Inter, 6.2.95).

Students knew that low letter grades brought few rewards. Summer (Gr.6) explained that she did not receive a reward for her grade 6 report cards because she did not do well enough (Inter, 5.5.95). She recalled the time in grade 5 when she received a C+ on her report card and her parents took her out for dinner. Summer pointed out that she no longer cared if she was unable to celebrate her report card.

Greg (Gr.5) said that he felt left out in terms of his peer group when he did not receive letter grades in Grade 4 (Inter, 5.16.95).

Greg: ...like last year I felt a bit left out because everybody else was getting letter grades.
Not really left out but like, I don't know I just felt weird. All my other friends got letter grades....

Martin (Gr.5) described his feelings of isolation when he received a low letter grade.

Martin: ...I like letter grades but I don't like it when I get too low ones....because I feel like the only person in the class who's not doing that really well in that subject....it makes me feel like I'm all by myself down low and everybody else us up high. (Inter, 5.16.95)

David (Gr.6) pointed out that a student's relationship with a teacher often influenced the marks he or she received.

David: Another reason for marks I would say is the teacher's relation with the student. If the teacher likes the student then the student's going to get better marks, but if the teacher doesn't like the student's style, just the student himself, then that's going to affect big-time on the marks....if the teacher likes you, you might not get a better mark but you'll get a more deserving mark, might get a better mark. If the teacher doesn't like you you'll probably get a less deserving mark....a teacher needs to like you if you want to do well. (Inter, 6.2.95)

Students explained that letter grades affected identity and self-esteem. Casey (Gr.5) said that she learned more about herself with each report card.

Casey: It (report card) tells you what kind of person you are...and it also tells yourself who you are....like you know yourself as yourself but you don't really know who you are....you don't know who you are in the sense that you don't know what you're really going to do, cause if you
really are yourself you can tell what you're going
to do in grade 12 like, you can say what school
you're going to go to, and you can even be
positive, but if you don't know yourself you can't
really make that decision...(the report card)
tells you more about yourself every time you get
one. (Inter, 6.22.95)

Casey added that letter grades told her how good she
was as a person (Inter, 6.22.95).

Matt (Gr.5) was more independent in that, although
the report card reflected teachers' perceptions of him
as an individual, he had his own opinions.

Matt: ...that's (report card) actually the
teacher's opinion of me but I have my opinion of
me....I think I'm a pretty bright kid....(Inter, 4.27.95)

David (Gr.6) was convinced that letter grades
should neither guide his reasons for learning nor shape
his identity.

David: You should tell you who you are and you
should know who you are, not what the little
letters on this little sheet of paper say. Those
shouldn't say anything 'cause if I drew a shape
that looks like an A rather than a shape that
looks like an F, that's just a drawing, just a
little thing on a sheet of paper.....the big
picture (report card)....I think it's pretty ugly.
It's not one of my favourite pictures....I
wouldn't buy it for a million dollars if it were
in some museum....It's sort of a false goal. It
should be to work well should be your goal, not
getting A's.....the big picture isn't always the
most important picture. (Inter, 6.2.95)

David argued against a view of education wherein
learning was exchanged for grades and his identity.

**Effort.** Students explained that, regardless of outcome, effort should be a criterion for measuring performance. Moreover, achievement should be interpreted according to the degree of effort applied to a task by an individual.

Casey suggested that there should be a place beside a letter grade on the report card to indicate how much effort was put into tasks. However, she thought that teachers' judgements of effort should not influence letter grades.

Casey (Gr.5): ...sometimes you might try really hard but you won't get any credit for it....(Inter, 6.22.95)

In contrast to Casey's views, Buffy explained that letter grades should reflect effort. Furthermore, merit was constitutive of effort.

Buffy (Gr.5): ...maybe...they can't draw, so maybe the teacher should see like that people try to draw but they can't...so like that's kind of one of their problems inside one of their disabilities, so the teacher should just give them a normal mark, like 3 out of 4....(Inter, 6.5.95)

Tricia (Gr.5) wanted a mark for effort to accompany grades because she wanted to know if she could improve her marks by trying harder (Inter, 6.6.95). Without a mark she was unable to determine if
she could improve (Inter, 6.2.95).

Although most students wanted Anne to take effort into consideration on their report cards, there was no agreement among students regarding how effort should be used in relation to letter grades. Some students wanted indexes of effort computed into letter grades. Other students wanted considerations of effort kept separate from letter grades.

Mediational Means

Tests. Students liked tests when they received high marks. They did not like tests when they received low marks. Children associated stress and anxiety with tests whether or not they did well. David described a test as "a big nerve-wracking thing" even when he knew that he would do well (Inter, 5.16.95, 5.8.95; Inter, 4.28.95; Inter, Bob, 5.5.95; Inter, Martin, 5.16.95; Inter, 6.2.95).

For Jeremy, a test served the sole purpose of generating marks for conversion to letter grades. He personified tests by saying that a test made up a letter grade.

Jeremy (Gr.5): A test is you have to prove how good you are to get a (letter) grade...It is an important part of your grade, it makes up your grade.... (Inter, 5.6. 95)
Bob thought that low test marks led to low letter grades.

Bob (Gr.6): ...if you get bad marks on your tests that will reflect on your report card. (Inter, 5.5.95)

Students understood tests as activities that generated marks in exchange for what they knew (Inter, 4.28.95; Martin 5.16.95; Brett 5.16.95; 6.6.95; 6.23.95).

Rick: We have a whole bunch of stuff before it (test) to work on, and then we study before it, and then we get the test and it tells us what to do. I mean it counts for our mark. (Inter, 6.2.95)

Students associated stress and anxiety with tests because they had limited prior knowledge of specific content. In contrast, they liked pretests because they knew what to study for a test. Within this context, they had more control over an activity and received higher marks (FN, 3.1.95; FN, 3.6.95).

Jeremy (Gr.5): ...if you're really bad at it and then you take a pretest and find out what you need to work on...then you correct it, and you could get it right in the test. (Inter, 5.8.95)

Martin (Gr.5): ...pretests help me a lot because then I can find out like what kind of stuff I will have to do and what kind of stuff I don't have to do so I can study that right thing and then I can like do pretty well on it. (Inter, 5.16.95)

Bob worried that he might fail a test by getting
low marks (Inter, 5.5.95). Meghan (Gr. 6) said that she was frightened by test schedules that did not allow her sufficient time to study (Inter, 4.28.95):

...when (the teacher) says, 'Okay we're having a test on math.' (it's) just shocking and then you just get scared....Like when it's a test you think 'Oh no, what if my, I get a bad mark and then it's recorded and everything!'...Cause Mrs. Wade has everything recorded and it also helps Mrs. Wade cause then she can look back and see your marks and say like you got like 16% on a social studies test, then she can tell that you haven't been doing your homework or anything like that.... (Inter, 4.28.95)

Matt (Gr.5) said that he did not receive good marks on tests because they made him nervous (Inter, 4.27.95). Due to stress and anxiety he experienced memory loss, increased heart rate, and tremors.

Matt: ...I do badly on tests because I just get really nervous...when I'm doing my regular work I'm getting in, for math for instance I'm doing two or three wrong now and then, but I normally get perfect. And it bothers me that when I'm doing tests I get like ten wrong....They're (tests) not harder and they're not easier. They're just like regular work but I get nervous....I don't think I get nervous but I am.

R: What are some of the things you feel when you're nervous?
Matt: Boom, boom, boom (says while hitting chest).
R: Your heart?
Matt: Yeh. My hands start to shake while I'm writing, my knees are shaking like a motor (makes sound of a motor). I'm just going like this, trying to be calm and I'm always like...'Oh, what is this?' and I know it's like '8+6, oh-oh-oh what is it? I can't remember, I can't remember! Oh, 14!' Finally, after that long! (Inter, 4.27.95)
Summer (Gr.6) and Bob (Gr.6) commented that they experienced memory loss before a test due to anxiety (Inter, 5.5.95; 5.5.95). Rick (Gr.6) was nervous before tests because he might not be able to do the questions, and low test marks might result in poor letter grades on his report card (Inter, 6.2.95).

David (Gr.6) argued that teachers did not need to test him to determine what he knew. He thought that his daily work should be sufficient evidence of what he knew. He explained that tests had limited value because they created stress, and did not necessarily reinforce important information and skills.

David: I don't think I should be tested. The work I've done should show it....it's necessary to be taught. It's not necessary to get worried about what you've been taught and to quickly sweat a lot or just write a lot about what you've been taught....no matter what in twenty years I'll probably forget the things that aren't, won't be important to me....so you remember what's important always and you don't remember what's not. So no matter what's on the test you're still only going to remember what you think you need to remember....you remember what you like, you remember what you need to know. (Inter, 6.2.95)

David saw test anxiety as an unnecessary expenditure of his energy.

Student self-evaluation. Occasionally, Anne invited students to mark their own work. Students
assigned numerical scores and written comments for creative writing assignments and research projects (Doc, 2.17.95; Doc, RP-95). They also completed self-evaluation packages regarding their report cards (Doc, RCR-95). Anne saw self-evaluation activities as means by which students reflected upon their learning, and participated in classroom assessment-evaluation practices. Students saw self-evaluation activities as means by which they communicated interpretations of their achievement to Anne and learned how to get higher marks.

Meghan (Gr.6) liked self-evaluation activities because they offered her opportunities to tell Anne what she thought of her work. She recalled that Anne gave her a 4 on an activity that Meghan believed was only worth a 3. Meghan wrote on the self-evaluation sheet that she had completed the activity "at the last minute" (Inter, 4.28.95).

Tricia (Gr.5) liked self-evaluation activities because they offered her a way to compare how she judged her assignments with Anne's judgements. In this way she saw how to improve her work and get higher marks. Tricia (Gr.5) also liked telling Anne in
writing what she thought of her work.

Tricia: ...then the teacher and I can see what I really think of how I've done....you can sort of compare how you think you've done to how the teacher thinks you've done and how the teacher wants you to do....Then you can know that you can do better....if you're not doing too well....if you're not trying your hardest. (Inter, 6.6.95)

Jeremy (Gr.5) regarded self-evaluation as a waste of time. He said that if his mark did not have the power to change the teacher's mark, then there was no point to his participation.

Jeremy: ...it doesn't matter what you think. It's what the teacher thinks so I don't think there's really a point to it....You're just wasting pen ink....like if she gives you a 1 and you put a 3, she's not going to change it to a 2. She's going to put what she thinks....(Inter, 5.8.95)

Jeremy's comments provided a stark illustration of the importance he placed on the exchange value of assessment-evaluation activities for marks and grades.

Similar to Jeremy's concerns, Cassidy (Gr.5) did not see the point of self-evaluation because she was not the one who made up her report card (Inter, 6.23.95). Matt (Gr.5) pointed out that he might receive a higher mark only if Anne valued his opinion.

Matt: ...she might find that your opinion is just not as good as hers but because, well she was taught to do this kind of thing. We are taught to work. (Inter, 4.27.95)
David (Gr. 6) suggested that it might not be appropriate for him to share interpretations of his performance with teachers. He explained that a student should be able to know how he or she was doing, but should not necessarily be required to share this information with a teacher. David focused on the use value of self-evaluation for his education to the exclusion of teachers.

David: You should be able to tell how well you're doing but I don't know if you should have to say how well you're doing. You should tell who you want to tell and who you feel deserves to know how you feel, so your teacher may not be that person. It may be your mother... it can be anyone. (Inter, 6.2.95)

Students saw self-evaluation activities as means by which they learned how to get higher grades, and communicated interpretations of their achievement to Anne. Although David saw the value of sharing self-evaluations, he did not think that he should have to share this information with teachers.

Overall, students saw self-evaluation activities as valuable because they served as means by which they learned how to receive higher marks. They also used the practice to communicate views concerning their work to Anne. Occasionally, a student met with Anne to
negotiate her or his marks for a particular activity.

**Criterion-referenced evaluation.** Students liked activities associated with Anne's understanding of criterion-referenced evaluation because they involved means for increasing access to higher marks and grades. They saw criterion-referenced evaluation activities as occasions when Anne involved them in setting performance criteria. From these activities, students learned that they received higher marks when they understood how Anne marked their work.

Students increased their access to high marks by negotiating performance criteria with Anne. Within this context, they clarified Anne's interpretation of criteria, changed criteria, and even rejected her interpretations.

Students explained that they received better marks when they understood performance criteria before engaging in a task.

Nicole (Gr.6): ...it (CRE) makes me get better marks because you know what kind of stuff you have to do....(Inter, 5.30.95)

Jeremy (Gr.5): ...if you cover every single one of those (criteria) you'll get a good mark. (Inter, 5.8.95)

Martin (Gr.5): I can find out what I need to do to get a higher mark. (Inter, 5.16.95)
Cassidy (Gr.5): ...it (CRE) tells you what kind of qualities you need to get a high mark.... (Inter, 6.23.95)

Casey (Gr.5): ...if you see it (criteria) before, you know what you're going to be marked on so you know what you have to work on.... (Inter, 6.22.95)

Meghan (Gr.6): ...I like having the criteria saying what you need to do to get...a 4. (Inter, 4.28.95)

Matt (Gr.5): ...you know what you have to do to get a certain mark. (Inter, 4.27.95)

Tricia (Gr.5): ...you know what you're getting marked on. (Inter, 6.6.95)

Meghan (Gr.6): ...you know what you need to get an A. (Inter, 4.28.95)

Oscar (Gr.6) thought that even though criteria guided his learning, criteria represented marks. Therefore, he would likely get a better mark by adhering to criteria.

Oscar (Gr.6): I know exactly what to do. Like if it says 'put some metaphors and similes in the story' I'll put some in.... R - So, that's criteria. Oscar - Yeh basically....well if they tell you you have to do that in the work that's basically a criteria...but if it's like you should do that it's not exactly like for the mark but you'd probably get a better mark if you did it. (Inter, 5.9.95)

Meghan (Gr.6) pointed out that rigidly set criteria limited her learning.

...it improves on what they're telling you to do, but not improves on what else isn't on the
Mechanics of conversion. Some students were preoccupied with understanding how to interpret marking and report card grading systems. Equally important, they were concerned with the degree of convertibility of one system to another. Students wanted to know how numerical scores converted to letter grades when assignments were marked out of 4, 10, 20, and so on. They also wanted to know how numerical scores were weighted in terms of letter grades.

Meghan (Gr. 6) explained that she preferred letter grades to numbers because she could compare letter grades on her assignments with letter grades on her report card (Inter, 4.28.95). She was unsure how to convert the value of the numbers 1–4 to letter grades. She added that the scale of numbers 1 through 4 did not offer as broad a window of achievement, as the scale of letters A through E.

Meghan: ...I just like As, Bs, Cs or Ds cause then you can compare it with...like your report card. You can say you got an A in science and it was with your invention and your project. Like say I got a 4, well I don't...I know what a 3 is, like a B?....I know 4 is an A, that's what Mrs. Wade (Anne) said. But like a 1 is probably like fail. I don't know what it is....a 3 seems bad to me. Like I don't know...but if I have a B that seems good to me because B is like, on the report card
it says what an A means and a B, and a B means excellent work... and so it just seems more than a 3. I don't know why. Cause I guess underneath a B is C, C-, C+ and D and E. But under the 3 is only 2 and 1. That's what it seems like, there's not much. (Inter, 4.28.95).

Some confusion over converting numbers to grades, and vice-versa, was due to the inverse relationship that existed between the two systems. The rank-order of letters represented a linear hierarchy with a top-to-bottom progression. The rank-order of the numbers 4 through 1 was reversed. Although 4 represented the highest exchange value for criteria on work sheets, 4 had the lowest hierarchical value. Students were used to interpreting value according to hierarchies of numbers and letters.

I felt disturbed when Martin (Gr.5) told me that he did not want to receive a 1, because 1 on the work sheet was represented by a flower stripped of its environment (Doc, 1.20.95), (Appendix E).

Martin (Gr.5): that sheet where if you're (4 it's) complete with birds and clouds and the tree and grass, and then in three they took away the clouds and the bird, and then in two there was only a flower with grass, and on one there's only a flower with nothing.... I don't want a one because there's only a flower. (Inter, 5.16.95)

I had seen these work sheets on numerous occasions, and did not interpret their numbers and flowers in this
way. Now I wonder if the blank space around the little #1 flower is even air.

I gained a deeper appreciation of Martin's discomfort by reframing the flower symbol in terms of a growth metaphor. Scheffler (1960) explains that an analogous relationship exists between a growing child and a growing plant, a gardener and a teacher. Although growth is somewhat independent, it is fostered and hindered by human intervention.

Martin's comments reminded me that marking and grading symbols came in many forms and are interpreted by children in many ways. I wondered how other symbols like check marks, (un)happy faces, and animal stickers, affected children's self-esteem, interpersonal relationships, motivation, and achievement.

When I asked David (Gr.6) what he thought of criterion-referenced evaluation, he said that it depended upon who created the criteria. After a moment he added that students should not always create criteria. David suggested that a student should be marked on her or his ability to do a task, and not against the performance of other students.

...there's no really criteria, for each student is different, each student should be marked on par
with their abilities, and that's it, that's flat, that's the rate, that all it should be...and not against criteria, not against other kids. ability's the only thing you have. You don't have really ah what you should do, you have what you do. (Inter, 6.2.95)

David's comments point to the value of self-referencing for descriptions of individual performance.

Overall, students liked criterion-referenced evaluation because activities involved means by which they learned more and received higher marks. Anne liked criterion-referenced evaluation because she saw it as an approach that supported learning and improved achievement.

Student-centered conferences and portfolios.

Students talked about their portfolios of work within the context of student-centred conferences. One hour of class time was scheduled for these conferences. Anne also held evening conferences. Parents were invited to visit the classroom to review their child's portfolio of work. Each student was expected to show samples of work in the portfolio to her or his parent(s).

For student-centred conferences on March 29, 1995, work samples involved stories, poetry, mathematics and French tests, science projects, artwork, chapter
question responses for social studies, work sheets, report card reflections, and special work (FN, 3.28.95, 3.29.95; Doc, 3.10.95; 3.29.95; 3.29.95a). Report card reflections involved students' perceptions of their progress in each subject, goals for improvement, and plans for achieving their goals (FN, 3.28.95; Doc, 3.10.95), (Appendix F).

Special work in a student's portfolio was flagged with coloured inserts. Anne placed a selection of the inserts on the rug at the front of the classroom. Students chose inserts to identify special work in their portfolios (FN, 3.27.95; Doc, 3.29.95).

Two days prior to the conferences, students prepared their portfolios according to the guidelines that Anne distributed (FN, 3.27.95; Doc 3.29.95a). Students referred to the agenda in the portfolio guidelines to conduct conferences with parents.

Some students saw student-centred conferences as means by which they used work samples to explain letter grades to parents.

Meghan (Gr.6): The student-led conference like backs up the letter grade. It shows all your work in that subject and everything. (Inter, 4.28.95)

Tricia (Gr.5): ...you can show the parents what you've done and not the teacher...I think the
child should be able to come because it's sort of not a secret what you've done....then your parents can understand why you got the mark you did.... Then they can match up your work to your mark. (Inter, 6.6.95)

Buffy (Gr.6) commented that the conference provided a means for telling parents her side of the story about marks.

Buffy (Gr.6): ...you can show your parents what you did and how you want to show it to them...and so you can talk to them about your work instead of the teacher....you can talk to them about your marks and how you feel, and with the teacher she just tells you how good you are, and then your parents know what you think. Cause then they know if we tried our hardest. If the teacher said that we didn't try our hardest but if we said we tried our hardest then they know that we tried our hardest....(Inter, 6.5.95)

Matt used the student-centred conference as a means to acquire praise.

Matt (Gr.5): ...I get to show the work that I did instead of Mrs. Wade showing the work that I did. It kind of seems like she's taking the credit for it....(Inter, 4.27.95)

Like Matt, Brett (Gr.6) enjoyed sharing special work in his portfolio with parents (Inter, 5.16.95). Brett pointed out that he grew uncomfortable when his parents could not find work samples in his portfolio. He planned to be more organized for future conferences.

When students were not doing well in school they did not like student-centred conferences. Conferences
did not provide a means for getting praise from parents when students received low marks and grades.

Nicole (Gr.6): ...if I'm doing well I like it (student-led conference), but if I didn't do so well on that sheet or whatever it was, I won't show them. I'll just tell them about it and then skip it. It won't be so bad....I'll probably just not put it in (the portfolio)......I would rather my mom than my dad (at the conference)....She makes me feel more proud of myself than my dad. (Inter, 5.30.95)

Martin (Gr.5): ...I don't really like them because if I do really bad work or something them my dad sees it....(Inter, 5.16.95)

Rick (Gr.6): ...it depends if you've done well or not. Cause if you haven't done well your parents will be mad at you....(Inter, 6.2.95)

Jeremy (Gr.5) pointed out that student-centred conferences were difficult for children of divorced parents, and children who did not have comfortable relationships with parents (Inter, 5.8.95). He added that the classroom might not be an appropriate place for some children to communicate with parents. Jeremy commented that student-centred conferences might negatively affect a child's relationship with parents.

Jeremy: I don't think student-led conferences are a good idea cause if your relationship with your parent is not that great, like say you've got a step-dad or maybe an uncle or a step-mom...you don't feel comfortable speaking eye-to-eye to, maybe you just wouldn't feel comfortable talking to them, and they wouldn't appreciate your work as much....I would rather take my folder home with
all my work, sit down in the living room, start a fire, get yourself comfortable and make a cup of tea or something, sit in your chair, invite your dad in and or your parents in, and tell them that you have some work from school to show them. And that would make it easier for both of you to concentrate cause you're relaxed and not alone in a big classroom instead of just being at home where you're comfortable. And it would be more easy for the student to talk about their work. (Inter, 5.8.95)

...Sometimes if you did some bad stuff and you put it in there (portfolio) it might affect your life with your dad or mom or something....Because you have this work and they thought you were doing really well on this, then they find out you did this one sheet really badly, and I guess it would make things go down hill....they wouldn't be as excited about coming to your student-led conferences again, and they won't be as interested in your school work....

...it's hard to decide which one (parent) you want to come....I could invite my step-mom and my dad, or my step-dad and my mom....It would be easier to talk about it at home at two different times. (Inter, 5.8.95)

David (Gr.6) had reservations about student-centred conferences for several reasons (Inter, 6.2.95). He explained that because he had not witnessed a parent-teacher conference, he was unable to judge if a student-centred conference was a better approach to communicating information about his progress to parents. Furthermore, he was unable to determine if the information he shared with his parents was accurate because his interpretations were mediated by his perceptions. Therefore, David decided that his
parents might learn more about his progress from the teacher. After a moment of hesitation, he added that he could not know if his parents would learn more from the teacher than himself because he could not speak for his parents.

David went on to explain that student-centred conferences were negative experiences for students who did not do well at school, and for students who were embarrassed to say that they had done well. He pointed out that some students might over-rate themselves.

David: ...I don't know what's better, the student-led conferences or when the teacher shows your work. I've never been to, I've never seen a teacher do a conference. I don't know if what I do is accurate. Maybe what I do might be what I want, so I don't think I really like it cause I don't know if it really shows me. It shows what I perceive of me. Or maybe what I perceive of me is what is me?
...It depends on what is you, that is who you are and how you perceive yourself. If you think you're this almighty god who knows everything, then my guess is you're going to totally over-rate yourself when you do a student-led conference....
...it's also a big question is whether you like doing it or not, if you find it embarrassing saying, 'I did this aren't I great?'
...for some people it makes them feel worse because say uh they gotta show, 'oh I did terrible.' (Inter, 6.2.95)

Overall David was sceptical of the value of student-centred conferences in relation to his education. Even though he was an A student, David
attended only some.

Most students saw student-centred conferences as means by which relationships with parents were mediated. When work samples in the portfolio exhibited low marks, some relationships were negatively affected. Some parents of students who received low marks were very supportive. When marks and grades met parents' expectations, relationships were positively affected. Students sought approval, praise, and respect from parents. Many feared disapproval from parents when they received low marks and grades.

Some contradictions were apparent between children's interpretations of student-centred conferences and Ministry policy. The policy document that focused on student-centred conferences explained that conferences provided students with opportunities to collaboratively explore and reflect upon their learning with parents (British Columbia Ministry of Education, 1994). Many students in the study pointed out that low marks did not foster collaboration between children and parents.

The document also described student-centred conferences as activities wherein "open, honest
dialogue is encouraged between teachers, students, and parents. Each participant benefits from the process" (British Columbia Ministry of Education, 1994, p.2). Jeremy's comments serve as a reminder that some children did not have comfortable relationships with parents wherein "open, honest dialogue" was feasible. In addition, as Jeremy pointed out, student-centred conferences were not designed to accommodate family structures that did not follow a traditional nuclear model of one mother and one father. From my standpoint, "open, honest dialogue" can be angry dialogue. Children did not want parents to be angry with them.

Conclusion

Within the contexts of Wertsch's (1991; 1995) concept of "mediational means," and Bourdieu's (1990) concept of "capital," I interpreted most students' perceptions of assessment-evaluation practices as means by which marks and grades were generated and differentially distributed for exchange on competitive school and home markets.

Many students reasoned that negotiating performance and evaluative criteria for activities
provided increased access to learning and high marks and grades. All of Anne's students were enthusiastic about criterion-referenced evaluation activities. In particular, they liked knowing what they needed to do to do well in an activity.

When students learned that marks and grades were mediated by performance and evaluative criteria, they wanted more control over criteria. Within the contexts of Wertsch's (1991; 1995) and Bourdieu's (1990) theories, I understood control over criteria as control over the authority of assessment-evaluation practices to mediate access to capital. Wertsch's (1995) story of how the gradual change from wooden to fibreglass poles increased vaulting performance, is a useful metaphor for understanding how learning and achievement increased when students participated in setting criteria.

Some students' efforts to access high marks and grades were complicated by the interface of their personal histories with assessment-evaluation practices. Bourdieu (1985e) explains that an individual's "habitus" in terms of dispositions, mannerisms, preferences, and world-view, is closely
linked to their practice and acquisition of capital. Students with histories of not doing well at school possessed little or no assessment-evaluation capital in the form of high marks and grades for exchange to social and other forms of capital. When habitus weakened by low marks and grades contributed to ineffective use of mediational means resulting in substandard performance and insufficient exchange capital. Greater control of assessment-evaluation practices by these students was central to their access to identity and well-being.

Students quickly converted capital from one form to another. At school, they quietly exchanged institutionalized capital in the form of marks and grades to social capital in terms of peer group acceptance, approval, and prestige. At home, they converted marks and grades to cultural, economic, and social capital in the form of clothing, food, compact discs, school supplies, money, and praise. It was not surprising to me that students wanted more control over assessment-evaluation practices when they thought that some practices gave them more control over their lives than did others.
Anne appeared to be caught in the middle between the Ministry of Education's expectations and parents' wishes regarding assessment-evaluation. Students were caught between Anne's and parents' expectations. Anne understood assessment-evaluation practices as means for monitoring student progress, determining achievement, and developing metacognitive skills. Many students saw assessment-evaluation practices as means by which marks and grades were generated. In the next chapter, I look at parents' expectations and interpretations regarding assessment-evaluation.
CHAPTER 7

We were all encouraged to be performing seals. I was a great performing seal and I think that's... a tragedy (Dawn, Inter, 6.6.95).

THE PARENTS

Parents And Education Policy

According to government officials and media, parents were prime actors in shaping B.C. education policy. In the Vancouver Sun on November 17, 1993, Premier Harcourt was reported as saying that parents and other members of the public worried that schools were not adequately preparing students for the future. People wanted greater emphasis on basic skills and clearer standards in terms of how achievement was measured and reported.

Many of the concerns that I (Mr. Harcourt) heard were about the Year 2000 program. British Columbians want to make sure that the basics are being taught...and that students' progress is being accurately measured and reported. (p.B1)

Art Charbonneau, Minister of Education, was quoted as saying,

We intend that parents get the information they need to determine for themselves how their child is doing. (p.B7)

The NDP government decided that a return to letter grades from grades 4 through 7 would facilitate
parental understanding of a child's achievement at school.

When I searched 1992 *Vancouver Sun* and *Globe and Mail* publications for evidence of parents' concerns prior to the policy change, I found the following statement made by Liz Shepherd, a mother of children in Vancouver elementary and high schools.

*It's a lovely system for nurturing and encouraging kids. It's not real life.* (Vancouver Sun, December 30, 1992, p.A5)

Mrs. Shepherd was concerned that a child-centred approach to learning in the primary program discouraged competition. In the same publication, Ian Shepherd, chairman of the parent-school council at Emily Carr Elementary School, explained that he was not convinced that proposed changes from Grades 4 to 12 would result in higher standards and stronger student performance. In addition, Keith Gray, a member of the BC Business Council commented that although schools' efforts to link with business were commendable, competition in schools was lacking.

*What we'd like to see is an awareness that there is competition out there* (p.A5)

In my exploration of parents' concerns after the policy change, I discovered a 1995 *Vancouver Sun*
article that represented the standpoint of Sylvia Dyck, president of the B.C. Confederation of Parent Advisory Councils. Ms. Dyck explained that schools should be teaching basic skills in terms of reading, mathematics, and writing, and socialization skills (Vancouver Sun, September 25, 1995, p.A4). Across from Ms. Dyck's comments was a list of criteria in terms of what employers looked for in potential employees. The list was developed from information generated by a survey of members of the Business Council of B.C. in May, 1995. Criteria were recorded in descending order of importance and included,


Ms. Dyck's emphasis on basic skills contradicted employers' lack of emphasis on the criteria, "well-educated." Employers did not equate "well educated" with criteria 1 through 13. Moreover, "well-educated" was classified in terms of something other than
"intelligent." If reading, writing, and arithmetic contributed to being "well educated," and employers equated "well educated" with low exchange value in terms of employment, then basic skills carried low exchange value on competitive job markets. What counted as education for parent groups was not reflected by employers' priorities; unless Ms. Dyck's interpretation of socialization skills included the other 14 criteria on the list.

Parents At Westheights School

Controversy regarding achievement was evident among Westheights parents. Like the children, parents in this study interpreted classroom assessment-evaluation practices as means by which marks and letter grades were generated. They thought that high marks and grades were important because they could be exchanged for access to high school, college, university, and the workplace. In some cases, they thought that marks and grades were important determinants of social standing.

Parents were relatively unaware of Anne's classroom assessment-evaluation practices. When I asked about their understanding of daily assignments,
tests, criterion-referenced evaluation, and projects, they said that they learned how their children were doing at school from report cards, student-led conferences, and homework. However, parents did not value these sources of information equally.

The report card was identified by parents as their most important source of information for interpreting how a child was doing at school. Some parents who read the report card insert recalled little of what was written. In this insert, Anne described how she used her interpretation of criterion-referenced evaluation in the classroom. She also provided a detailed overview of content and skills that were covered in each subject area during the term.

Most parents were preoccupied with letter grades. They equated high letter grades with high achievement, and low letter grades with low achievement. They wanted letter grades on the report card because they had a history of convertibility to employment.

Some parents saw student-centered conferences as opportunities to see how marks and written comments matched up with some work samples (Inter, 5.30.95; 6.5.95). Conferences also represented opportunities
for several parents to praise children who received good marks and grades (Inter, 5.24.95; 5.25.95; 5.30.95; 6.8.95; 6.10.95). One parent who was regularly involved in her son's work thought that student-centred conferences told her little, if anything, new about her child's progress (Inter, 4.27.95).

Homework was not a concern for parents of children who received good letter grades. Parents of children who did not receive good letter grades said that homework supported their child's learning because extra work improved skills. Overall, parents did not see homework as an important source of information concerning a child's achievement.

In this chapter I examine parents' interpretations of classroom assessment-evaluation practices as means by which marks and grades were generated for exchange on the job market. Following this discussion, I explore how parents understood letter grades as their most important source of information concerning a child's achievement.

Exchange Value

Training for the workplace. Similar to children's
interpretations, parents emphasized the exchange value of marks and grades. "Mary," a Chinese woman, explained that high grades were exchanged for privilege in Canadian society wherein different races bore different exchange values on competitive markets (Inter, 6.5.95). She pointed out that being "white" had a higher exchange value in the Canadian job market than did being Chinese.

Mary explained that she wanted her son's schooling to train "him to be a white person, and get a job" (Inter, 6.5.95). Her son, Martin, was the little boy who did not want a score of 1 on his work sheet, because 1 was associated with a flower stripped of its environment. As I listened to Mary, I wondered if Martin saw himself as the flower; stripped of his culture by Canadian assessment-evaluation practices?

"Nancy" said that she wanted to know where her son stood compared to the rest of the students in the class. She explained that parents in Asia were told how a child ranked in relation to her or his peers. Nancy added that it was a matter of "face" for parents that children did well at school. Parents exchanged grades for social standing (Inter, ?)
"Marge" commented that, due to rising unemployment, schools should prepare children to become self-employed adults. She wanted assessment-evaluation practices to prepare her children to compete for the primary stake in the economy; money.

Marge: ...this is what society is geared for, how much money, how successful can you be in making a business run. Because, let's say there's really no jobs out there...I mean there isn't enough work for people. You have to create...your own jobs...start a business....sure learning about the world is great....but how about learning about the economics....(Inter, 5.25.95)

Similarly, "Brooke" pointed out that the workplace was based upon competitive achievement. She wanted the school's assessment-evaluation practices to tell her where her children's performance ranked in relation to their peers. She wanted to know how effectively they competed with their peers for marks and grades.

Brooke: I need to know how he's doing in his own progress....But I also need to know where he fits with everybody else....She (Anne) might give it to me if I said, 'I want to know...is he 15 out of 21, is he 20 out of 21?' ....When you're going competing for a job you're being compared with the other applicants so, it's there....when (my father) ran his law firm (he) would say, 'well we hire students who are at the top of the class.' (Inter, 4.27.95)

Brooke added that children compared marks and grades with peers regardless of teachers'
interpretations of assessment-evaluation practices. She went on to explain that a teacher might see assessment-evaluation practices as activities that encouraged learning, but these practices fostered competition among children.

...we all know the kids are comparing themselves to everybody else....There isn't a child in the universe who doesn't realize that, and we're fooling ourselves if we think we're...making it fuzzy enough so they're not doing that....So you can call it any name you want....the Bell Curve, you can call it norm, you can call it criterion based, kids are comparing themselves....It might be more subtle, it might be politically incorrect but they're still doing it and everybody's doing it. (Inter, 4.27.95)

Similar to Brooke, "Betty" linked assessment-evaluation practices with the workplace. Betty explained that students learned to exchange work for grades in preparation for a workplace wherein adults exchanged work for money.

Betty: ...school in a way prepares you for the real work later on....you're there to accomplish a certain number of things in a certain period of time. And I see that as a job the kids have to do. I go to work and make money for us to live and they go to school and learn, that's their job.....(in the work place) you're not going to get letter grades but you're going to get X in salary. (Inter, 6.11.95)

Kate explained that a capitalist society was rationalized by acts of comparison among people. She
described how letter grades operated as sources of comparison in the following manner.

Kate: ...being that we operate in a capitalistic society you do need to have some comparators ....it's (letter grade) a comparator and it tells you basically within a range how you're doing compared to other people....(Inter, 6.12.95)

Interpreting Letter Grades

Parents explained that letter grades told them how their child ranked in relation to curricular expectations and/or class performance. The importance that parents placed on letter grades recalls the Minister of Education's claim that letter grades gave parents "...a better understanding of how their child...(was) progressing" (British Columbia Ministry of Education, NR29-94, p.23). Mr. Charbonneau's claim implied that parents shared a common understanding of letter grades. Although parents in the study placed importance on letter grades, their interpretations of letter grades differed.

What's in a letter grade? "Sue" and "Dan" took-for-granted that their interpretations of letter grades were similar. After many years of sharing their children's report cards, they were surprised to learn that their interpretations differed.
Dan: I suppose a letter grade...at least in theory they supposedly you know rank the students in relative order of merit to one another.

Sue: Oh not to me! It means that in relation to the stuff he's supposed to be learning he's doing it really well. I don't think of it in relation to the other kids in the class. (Inter, 6.10.96)

Although Dan and Sue understood letter grades as symbols of comparison that represented their son's achievement, they did not agree upon the frame of reference for comparison. Dan assumed that letter grades represented teachers' judgements of Miles' performance compared with the performance of his classmates. Sue assumed that letter grades represented teachers' judgements of her son's performance compared with expected learning outcomes of the curriculum.

Unlike Sue and Dan, "Wendy's" and "Blake's" interpretations of letter grades were similar. They assumed that letter grades told them how their son was doing at school in relation to a standard. However, Wendy was not sure if a letter grade ranked her son's performance in terms of curricular expectations and/or class performance. She wondered how teachers constructed letter grades.

Wendy: I guess the biggest thing that it (letter grades) tells us is how they (children) rank in comparison to what was expected. Perhaps also to
where the rest of the class was but I don't know how much they use this bell curve thing... (the letter grade) just tells you a comparison of the child with the ideal...but it doesn't say what the ideal is... (Inter, 6.19.95)

Unlike Wendy, Blake interpreted letter grades in terms of hierarchical value. Some children also used a hierarchical approach to creating meaning for letter grades.

Blake: Well in 30 seconds of looking at the report card you have a pretty good idea of whether your child is within reasonable performance parameters. If you see lots of D's and C's then your child's really not performing up to standard...if you see above well...your child's probably doing reasonably well in school. (6.19.96)

"Dawn" assumed she had an interpretation of letter grades until I asked her what they meant. Similar to Wendy, Dawn indicated that she was confused about the nature of comparison used to generate letter grades. In the following excerpt, Dawn struggles to construct an interpretation of letter grades.

Dawn: ...I grew up with it (letter grades). It's like a red light, a green light and a yellow light. I know what it means...it's a cultural, it's sort of a cross-cultural thing. It's like, I don't know, just a signal right? We all know what they mean and so that's easy....They tell me where on the strata his work is....I know what C+ work is....It's like midway right? On the scale, thermometer, or the range you know?....Well if cold is over here, and hot's over there, you know we all know where a C+ is right? It's...in the middle....I don't know, is it (letter grade)
measured against the other marks in the class?....
R: Take the average you mean, the class average?
Dawn: No, they have a certain, you know they, out of this many kids, a certain number of them are going to have As, and a certain number are going to have Bs? (Inter, 6.6.95)

Dawn used hierarchical metaphors to interpret letter grades. Her use of metaphor mediated her interpretation of letter grades. Equally important, her choice of metaphor was mediated by her experiences. Dawn concluded that letter grades told her where Jeremy ranked in relation to a comparative standard. However, she did not understand the nature of the standard teachers used to evaluate her son's performance.

"Barb" remarked that when letter grades were meaningful they told her how well Angela was doing at school in terms of expected learning outcomes. She added that she was not interested in how Angela's performance compared with the performance of her classmates. Barb pointed out that classroom practices that measured student performance in relation to curricular expectations could result in all students receiving As.

Barb: I don't care so much where my child is in relation to the other kids in the class but where she is in relation to what's expected for that grade level....It's not like one of these you know the top 5 in the class get A. I mean to me that's
not important...you could have a class where an entire class was over 80%...then everybody should be A....If that's what ABC means...because I'm not interested in anybody else in the class except my kid....I think that's what most parents are ...(Inter, 5.29.95).

In contrast to Barb's interpretation, "Grace" and "Kate" said that letter grades told them how their child was doing at school compared with the rest of the class (Inter, 5.25.95; 6.12.95). "Linda" and "Betty" also interpreted letter grades as indicators of their child's performance compared to class performance.

Linda: The only thing it (letter grade) tells me is if he's average or he's doing better than average....I think a letter grade is more like an average of what all the children are doing...doesn't seem to be really specific to what he's doing. (Inter, 6.6.95)

Betty: ...it's (letter grade) a measurement...it's the system that I have worked with...through all my school years through to university and college level where at the end of the day you're given a tag and you know where you fit within the hierarchy...I don't know whether human nature needs that or it's simply they're just accustomed to that process. (Inter, 6.11.95)

Parents explained that confusion over meaning of letter grades was exacerbated by inconsistencies among teachers' interpretations.

Wendy: ...maybe it's a false assumption that letter grades are going to be consistent. (Inter, 6.19.95)

Hope: Is an A in one class the same as an A in
another class? (Inter, 5.10.95)

Judy: ...each teacher has...her own style....It's not 100%....I mean one kid could be really good in one situation and be pretty lousy in another situation....It's not black and white....it's hard to have that consistency cause we're all human, we're not machines. (Inter, 5.25.95)

"Sue" commented that regardless of teachers' interpretations of letter grades, student performance was compared to some form of criteria. She added that discussion was needed to address whose norms mediated criteria.

Sue: ...you get into...does this particular indicator really relate to that particular outcome? Is it valid? Is it a reliable indicator? Is it this'or is it that? And then you think 'oh geez, what does it matter anyhow?'....Whose norms, what values? Do they cross culture, do they cross whatever? You know, I could get into the same arguments for both systems (norm-referenced and criterion-referenced evaluation) ....what are the values that these norms are based on? (Inter, 6.10.96)

Sue's comments recalled how Anne and students discussed and negotiated criteria to co-construct a performance norm for an activity wherein students made Dr. Seuss booklets. The norm involved beliefs and values of participants concerning what criteria should count as acceptable performance.

Anne interpreted criterion-referenced evaluation as valid because it was pedagogically sound. That is,
students learned while demonstrating their understanding of previous and new skills in a manner that was appropriate to their strengths, weaknesses, and interests. She interpreted the approach as reliable, or useful, because it made sense in relation to other classroom practices. Anne selected and constructed criteria that met her expectations regarding what and how she wanted students to learn.

When interpretations of letter grades turn on questions concerning whose norms are used in their construction, and people have different interpretations, then confusion or clarity on hierarchy is not surprising.

The price of letter grades. Although parents equated letter grades with exchange value in terms of employment, some pointed out that grades failed to credit dispositions and qualities important to the growth of community, citizenship, and collaborative work relationships.

For example, Brooke explained that As on a report card did not ensure that a student would become an effective employee.

...this is my problem with evaluation because we all know millions of people in the workplace who
have a fantastic letter grade but they aren't terribly effective....maybe because they can't communicate with people or maybe because they don't have a sense of humour. Or maybe cause they aren't team players or they can't show leadership. (Inter, 4.27.95)

Brooke's comments are interesting in light of what we learned from Anne, her colleagues, and Graham the school principal. Specifically, teachers used different practices to gather information and make judgements about student progress. In addition, they interpreted these practices differently. Consequently, an A from one teacher did not mean the same as an A from another teacher.

Brooke pointed out that letter grades did not credit a student who demonstrated integrity, loyalty, and consideration for other people. Moreover, low grades had a negative impact on self-esteem.

...the comments aren't part of the grade. The comments are in addition to, 'oh by the way and even though your kid is a C- kid, he's a nice boy'....I run self-esteem groups and one of the big problems with self-esteem is people come in and say, 'well I'm a really nice person, I've got integrity, I'm a loyal friend...but so what, I didn't get into grad school because...my undergraduate marks are 75, so where does all this being a nice person, being an honest person get me? So what!'....Our society isn't valuing any of those things. And the kids are saying 'well I might be a nice person...but...I'm not smart.' ....I've had a couple of different teachers over the years say to me about both boys, 'they're
really nice boys... I guess that doesn't... count for anything because we are really measuring academics (Inter, 6.12.95).

Brooke’s comments illustrate the power of exchange value to shape self-esteem, identity, and opportunities for participation in society. Smart people competed successfully for marks and grades. If students learned that being nice did not count in school, and that being smart equated with high grades, then what happened to a student with low marks who gave up on being nice? I remember teaching students who were failing school, were not nice, and were on probation.

Even though parents wanted their child to receive high grades, they also wanted their child to have a positive attitude towards school and learning. The down side for parents of equating assessment-evaluation practices with exchange value, was the negative impact that low grades had on self-esteem.

"Joan" commented that she had mixed feelings about letter grades when Casey became upset with a C+ she for Art (Inter, 5.16.96). "Liz" liked seeing letter grades on the report card, but she was concerned about their impact on Nicole's self-esteem.

Liz: ... you sort of pay your price for having letter grades because if she (Nicole) sees she's
only a C or C+, if she's lucky, she goes 'my report card isn't really that good.' So then you try to explain to her 'maybe you could get a B-, wouldn't that be great if you brought that up?'...when Nicole brought home her first report card...it wasn't a good report card...So I sort of went 'oh no, she's just going to be devastated because...her letter grades weren't very good....(Inter, 5.29.95)

Joan's remark to Nicole that a higher grade would make her feel better, is a graphic illustration of how a child learned to equate learning with exchange value; high grades converted to social capital in terms of self-esteem.

Brooke described how her son cried at night when he did not do well at school. Matt equated his low marks with being stupid.

Brooke: ...we still have our bad days, you didn't see him in September. Anne didn't know how he was crying every night, but I told her. But he's certainly not doing that now. He's feeling, he's not coming home every day saying, 'I'm the stupidest one in the class and I can't keep up,' which he was doing at the beginning of the year....back in September he had practically, he thought he couldn't read. He had lost his confidence to the point where he couldn't read. (Inter, 4.27.95)

Kate said that poor letter grades demotivated students. When they received low grades they stopped trying to achieve (Inter, 6.12.95). Dawn described the emotional instability created by fluctuations in a
Dawn explained that an emphasis on letter grades when she was a student influenced her to equate learning with performance (Inter, 6.6.95). She did not want Jeremy to experience her frustrations as a student.

Judy remarked that students who did well at school liked to compete for letter grades. However, she pointed out that letter grades might be discouraging for students who did not do well.

Judy: ...I think grading is important for kids. I think kids are competitive and they like to be the best in the class, or the second best or whatever. For the bright kids certainly. The lower level kids I think it may make them feel awful, always getting low....so you know it's a catch-22....I mean we want to encourage our students both the weaker ones and the stronger ones...everyone should be encouraged and should be feeling good about themselves by doing the best they can. So really, I don't know if we want to have letter grades or not. (Inter, 5.25.95)

"Gwen" explained that her daughter's confidence...
was undermined when she received letter grades less than a B (Inter, 6.8.95). She added that Buffy had grown to dislike subjects for which she received C's.

Conclusion

Parents felt that marks and grades could be used as exchange on competitive school and job markets. High marks and grades were important to parents because they had high exchange value for educational opportunity, academic credentials, employment, money, and social standing. Parents were used to letter grades with a history of convertibility to academic credentials, employment, and money.

The report card represented the most important source of information for parents because it was an official record of student achievement. Letter grades on the report card mediated access to promotion, academic credentials, and later employment. The report card had the authority or symbolic power to define, classify, and make legitimate, various classifications of achievement. For some parents, the report card was a form of capital that conferred identity on an individual, which in turn granted cultural and economic advantages. Several parents commented that low grades
negatively affected self-esteem and suggested that honesty, consideration, and collegiality should be reflected by letter grades.

Parents' interpretations of policy and classroom practices associated with assessment-evaluation were mediated by their experiences. They drew from experiences when they were students and from their work as employers and employees. Parents neither referred to report card definitions of letter grades as a frame of reference for their interpretations, nor the Ministry document, "Supporting Learning: Understanding and Assessing the Progress of Children in the Primary Program" (Doc, RC-1995; Doc, 1992). Grant issued copies of "Supporting Learning" to the school's parent population on several occasions (Inter, 5.18.95).

Even though parents had different interpretations of how letter grades were mediated by assessment-evaluation practices, they were unanimous in equating letter grades with exchange value. Even when parents recognized that low letter grades had a negative impact on student identity, they supported a system of education defined by competitive achievement. In this way I found that what Premier Harcourt said was true.
Parents wanted letter grades and interpreted them as exchange for employment and money.

In contrast to most parents' interpretations, Anne equated assessment-evaluation practices with use value in terms of supporting learning, improving achievement, and helping children learn how to learn. However, in line with Ministry policy she translated assessment-evaluation information into letter grades on report cards. Like most parents, the majority of children equated assessment-evaluation practices with exchange value. That is, assessment-evaluation practices generated marks and grades for exchange on competitive school and home reward markets.

Although I gained some insights into the policy change from the standpoints of administration, teachers, students, and parents, I learned that there was no consensus among participants.
CHAPTER 8

"To change the world, one has to change the ways of world-making" (Bourdieu, 1987b, p. 23).

CONCLUSION

This conclusion looks first at the meaning of findings from Anne's standpoint, and then within the context of Bourdieu's (1990) and Wertsch's (1991, 1995) theories. Following this discussion, I examine implications for theory, practice and policy, and further research.

My overall conclusion is that assessment-evaluation procedures and practices are socially situated in settings characterized by unequal relations of power among policy-makers, teachers, students, parents, and administration. A useful approach to assessment-evaluation would involve more equitable relations of power among these people. Greater communication and collaboration within, and across, interest groups may facilitate growth of shared understanding regarding what counts as useful assessment-evaluation policy and fair assessment-evaluation classroom practices. In particular, teachers need more time and support to explore,
discuss, and reflect upon assessment-evaluation policy and classroom practices within and across contexts of their work.

Anne's Reflections

This study was a collaborative project. Anne and I jointly collected and examined information regarding her classroom procedures and practices. Anne contributed to the writing of this thesis by editing her interview transcripts and drafts of Chapter 5, and by writing her reflections and some recommendations for the final chapter.

We agreed that mutual trust was central to our collaborative relationship and that mutual trust was built upon honesty, consideration, and mutual respect. It was not always easy for Anne and me to be honest. We did not want to hurt each other's feelings. I recall one afternoon when Anne told me that I was becoming a nuisance. I quickly apologized and got out of her way by cutting mural paper, washing paint bottles, and hanging pictures in the hall. Even though her comments were a blow to my ego, I reminded myself that it was report card time. Anne was overwhelmed with planning, teaching, collecting and marking
assignments, and writing up report cards. She did not have time to oversee my activities.

Later that same evening, I telephoned Anne and explained that I was feeling uncomfortable with our relationship because she appeared to be unhappy with my presence in the classroom. She replied that everything was fine and we would get to know each other with time. Anne apologized for giving me the impression that I was not welcome in her classroom.

Early in the data-collection phase of the study, Anne and I made a pact that we would not compromise our honesty with each other for anyone or anything. Challenges to the integrity of our relationship were unpredictable and ongoing. I grew to treat each day of the study as its last. Anne and I were unable to assume that successful completion of one day guaranteed successful completion of the next. The upkeep of our collaborative relationship continues to involve deeply satisfying work, good times, and the occasional surprise. Because of the importance of our collaborative relationship to this thesis, I begin this conclusion with Anne's reflections and reactions to the collaborative process as a whole (Doc, 4.12.97).
The thesis is useful to teachers, students, and parents because it points out that kids have a lot to offer about issues that affect them. And, they should have input into decisions made about educational issues that directly affect them. From this study, parents can gain an understanding of what assessment practices look like in the classroom, and some idea of how a child's marks and grades are constructed from those practices. Furthermore, parents and teachers can gain insights into what is important to children and how children feel about the assessment practices they participate in, and have imposed on them. Teachers seldom know in detail what happens in another teacher's classroom with regards to assessment. They can get another perspective on assessment issues from this study.

The Ministry of Education is removed from the classroom setting. This research will give people associated with the Ministry a sense of what assessment practices are used in a classroom, how they are used and why, and how assessment practices affect students and teachers. This research has affected my classroom practice by giving me insights into my assessment practices. I find it useful to know what students and parents think about my approach to assessment. The study has taught me that children have great insight into what affects them, and they are capable of sharing important information about assessment with adults. This means that their opinions are valuable. Teachers and parents, everyone involved in education, should ask for their opinions on assessment and schooling in general.

In my collaborative relationship with Karen I became more aware of my practice. We were both open to the idea of collaboration. We were open to sharing and neither one of us felt the need to be the person in control. Karen and I respected each other and valued one another's teaching experiences and strengths. Early on we agreed to be frank with each other about concerns. Our teaching styles, how we view children, our
philosophies of education were similar. Together we developed useful and effective assessment methods and units of study. I enjoyed working collaboratively, something that teachers rarely get to do. As I look back on it I can say it was a lot of fun. (Doc, 4.12.97)

Summary Of Findings

My interpretations of findings are located in the conclusions of chapters four through seven, and are framed by Wertsch's (1991; 1995) sociocultural theory and Bourdieu's (1990) theory of genetic structuralism.

Policy and assessment-evaluation. Policy-makers were concerned with getting the public on side. NDP government officials said that a change to criterion-referenced letter grades responded to parents' requests for clearer standards regarding achievement. In particular, a return to letter grades from grades 4 to 7 would make it easier for parents to understand how a child was doing at school.

Parents in this study wanted letter grades but did not know how they were constructed. Unlike Anne, parents did not assume that the meaning of a letter grade was framed by its construction. Anne described in detail her approach to creating letter grades in Chapter 5. Although many parents assumed that an A meant that a student's performance was above-average,
Sue, thought that a letter grade represented how well a student demonstrated understanding of curriculum criteria.

The term, criterion-referenced, was used in policy documents though its meaning was unclear. The relationship between criterion-referenced evaluation and assessment was difficult to determine. The relationship between criterion-referenced evaluation and criterion-referenced letter grades was equally difficult to understand. This lack of clarity generated confusion for Anne and the school principal.

Anne assumed that criterion-referenced evaluation and criterion-referenced letter grades were related because the phrase, criterion-referenced, was common to both terms. She assumed that criterion-referenced letter grades should be constructed from criterion-referenced evaluation activities wherein students negotiated performance criteria. Because Anne did not have enough time, nor the expertise in some areas like Art, to implement her understanding of criterion-referenced evaluation in all classroom activities, she assumed that her letter grades were not criterion-referenced. This worried her.
Learning and assessment-evaluation. Anne assumed that the purpose of assessment-evaluation was to support and promote learning. Anne thought that letter grades had a long history of ranking students, and that their history made it difficult for teachers, students, and parents to disassociate letter grades from ranking procedures.

Anne's understanding of, and responses to, the policy change were influenced by her personal history, teacher training, professional development, classroom experiences, colleagues, and graduate study. Her involvement with colleagues on a "Teacher-As-Researcher" committee and during professional development day workshops, led to her understanding of criterion-referenced evaluation as an instructional strategy. Anne developed and modified her interpretation of criterion-referenced evaluation in response to students' needs, strengths, and interests.

Consistent with my understanding of assessment-evaluation, Anne assumed that classroom procedures and practices associated with collecting and making judgements about student progress based upon curriculum criteria were interconnected. She referred to these
procedures and practices as assessment. This use of
the term, assessment, conflicted with policy
discussions that drew sharp distinctions between
assessment and evaluation. Although Anne did not adopt
the Ministry's uses of the terms, assessment and
evaluation, she regularly collected assessment
information regarding student progress, and evaluated
this information.

Anne's assessment-evaluation information involved
tests, projects, presentations, daily assignments,
assignments, homework, portfolios, observations, teacher-student
conferences, student self-evaluations, marks, a mark
book, written evaluative comments, and letter grades.
She made judgements about a student's progress by
comparing her or his assessment-evaluation information
with program objectives, past performance, and/or
present performance in relation to curriculum criteria
and expected learning outcomes. In this way, Anne
determined a student's strengths and where improvements
were needed. She assigned letter grades according to
how effectively a student met performance criteria
throughout a term. All of Anne's approaches to
assessment-evaluation including letter grades were
criterion-referenced because she based them on curriculum criteria and expected learning outcomes.

The policy change had a negative impact on Anne's working conditions. Because she was unable to involve students in negotiating criteria for activities from which she gathered information to create letter grades, Anne thought that she was not doing her job as well as she should. Consequently, she worked overtime to address the needs of students, colleagues, parents, administration, and to implement her understanding of the policy change in her classroom practice. From my standpoint, Anne was caught between poorly written education policy regarding assessment and evaluation, and her immediate need to understand and respond to the policy change in her classroom practice.

The Dr. Seuss activity is a very good example of Anne's integration of assessment-evaluation practices and instructional strategies within the context of a language arts lesson. Anne observed how well students understood performance criteria throughout their negotiation of criteria, and while they created Dr. Seuss booklets. After she collected booklets and criteria work sheets, Anne compared her observations
regarding student performance with students' self-evaluations of their performance. When assigning marks and writing comments for the activity, she took students' views into consideration. When Anne returned the booklets and criteria work sheets to students, she told them where to keep the work until it could be filed in their portfolios.

Grant understood the policy change as a shift in emphasis from norm-referenced evaluation to criterion-referenced evaluation. He saw the shift to criterion-referenced letter grades as the most critical change for teachers. Grant was understandably confused regarding uses of the terms, criterion-referenced evaluation and norm-referenced evaluation. He interpreted a practice as norm-referenced when norms of achievement were developed from a large population. However, he was not certain how large a norming population should be. Although Grant understood the Ministry's claim that norm-referenced procedures were not appropriate for classroom use, he could recall years when norm-referenced procedures were popular fare.

Anne and Grant attended some of the same School
Board workshops on criterion-referenced evaluation. Consistent with Anne's understanding of criterion-referenced evaluation, Grant interpreted a practice as criterion-referenced when teachers discussed and set performance criteria with students. Within this context, he understood a criterion as an index of performance that was collectively established by teachers and students.

Grant was not concerned that teachers on his staff constructed letter grades differently. He regarded diversity as an important quality of professional judgement. Overall, he thought that administration, teachers, and parents did not understand the policy change. He suggested that common understanding regarding the policy change must begin with teachers.

Kids, parents, and assessment-evaluation. Most students and parents were concerned with getting ahead. They understood and responded to classroom assessment-evaluation practices as means by which marks and grades were generated for conversion on competitive school, home, and employment markets. Many students thought that high marks and grades had high exchange value for clothes, food, money, CD's, friendship, positive
relationships with teachers and parents, identity, access to university, and perhaps jobs.

Within the contexts of Bourdieu's (1990) and Wertsch's (1991; 1995) theories, I understood students' comments as evidence that assessment-evaluation practices operated as mediational means that generated capital in the form of high marks and grades. From this standpoint, it is not surprising that students liked assessment-evaluation practices when they got good marks, and did not like those practices that resulted in low marks.

Most students thought that some classroom practices, like criterion-referenced evaluation, gave them greater access to high marks and grades than did others, like tests. These students liked Anne's approach to criterion-referenced evaluation because they thought they learned more and knew what they needed to do to get high marks. By collaborating with Anne to set criteria for an activity, they gained a clear understanding of performance requirements and how their work was evaluated. Within this context, students learned to address specific skills and content in a manner consistent with Anne's expectations. Many
children thought that they learned more, and achieved greater success in their activities, when performance and evaluative criteria were collaboratively constructed. It is not surprising that students grew to want more control over assessment-evaluation classroom practices.

Many students were adept at accumulating and trading marks and grades for friendship, praise, and goods and services. One student who was unable to accumulate grades, stopped trying to compete. Overall, students were caught between Anne's emphasis on learning in relation to assessment-evaluation practices, and parents' preoccupation with high marks and grades.

Most parents regarded the report card as their most important source of information because it represented the official record of student achievement. They liked letter grades because they were on the report card. These parents wanted letter grades because they thought that letter grades had high exchange value for educational opportunities, jobs, and money. However, there was no consensus among parents regarding how letter grades were determined. They were
relatively unaware of Anne's classroom assessment-evaluation practices, and had not read policy discussions regarding assessment and evaluation. Most parents with whom I spoke could not recall Anne's discussion of criterion-referenced evaluation in report card inserts.

Parents based their interpretations of letter grades on their experiences at school and in the workplace. Many saw school as a training ground for entry into the "real world"; a world of work defined by competitive achievement. These parents interpreted letter grades as symbols that compared a child's achievement with curriculum criteria and expected learning outcomes, or with the performance of other children in her or his class. Even though some parents expressed concern that low letter grades might have a negative impact on a child's self-esteem, they generally supported their use.

Parents ascribed more value to the report card than to classroom assessment-evaluation practices because the report card represented the official record of a student's achievement. Even though Anne's assessment-evaluation practices generated marks and
grades, these practices did not represent official records of student achievement.

Most parents focused on outcomes of assessment-evaluation practices. By overlooking Anne's assessment-evaluation practices, parents lacked awareness of performance criteria and procedures associated with marks and grades. Many parents reinforced a positive relationship between learning and exchange value by rewarding kids for high marks and grades. So, Premier Harcourt was right. Parents wanted letter grades. But the relevance of this observation to learning is questionable.

**Implications For Theory**

By working collaboratively with Anne, I grew to see criterion-referenced assessment-evaluation as a socially situated practice. Wertsch's (1991; 1995) and Bourdieu's (1990) philosophical and methodological standpoints proved to be useful frameworks for examining and interpreting assessment-evaluation as socially situated practice.

The present study extends an understanding of Wertsch's and Bourdieu's interpretations of the dialectical relationship between agency and structure.
through an approach to criterion-referenced assessment-evaluation as situated in a social context. Within this context, assessment-evaluation is socially situated by, and within peoples' interpretations of, and responses to, education policy, classroom assessment-evaluation procedures and practices, and cultural, historical, and institutional settings.

Many students have little or no control over assessment-evaluation practices in terms of selection, performance criteria, marking schedules, and contexts for use. Anne's interpretation of criterion-referenced evaluation in relation to some activities gave students greater control over performance and marking criteria than they had in relation to others.

Wertsch's (1991; 1995) observation that human agency is constrained by the very logic of its production, is not pessimistically reproductionist. That is, restraints on human agency can be changed and even replaced by others. On the one hand, Anne increased student access to achievement by sharing her power to set performance criteria with students. On the other hand, the approach was time-consuming and her ability to involve student participation was limited.
This limitation generated difficulties for her practice and led to stress.

Some adults in this study had more power than children to name what counted as achievement. Policymakers and school administrators had more power than teachers to decide what counted as achievement. The premier's comments regarding letter grades might imply that parents also had more power than teachers to define achievement. These tensions raise questions regarding the relationship between politics and professional expertise.

Contradiction, conflict, and uncertainty are emblematic of activity mediated by unequal relations of power. Theoretical constructions should explore the concept of power more thoroughly as it applies to assessment-evaluation within the contexts of Bourdieu's and Wertsch's theories. Their theories are useful because they address issues associated with power within and across disciplinary fields such as sociology, psychology, and education. In other words, more theoretical work should be done regarding assessment-evaluation. Equally important, theorists should locate and construct their work within and
across contexts of assessment-evaluation.

Writing about Bourdieu's (1990) and Wertsch's (1991; 1995) interpretations of dialectical relationships was very difficult. The left-to-right progression of written English forced me to organize my discussion of dialectical relationships in left-to-right bytes of developmental and sequential text. These limitations crippled my efforts to describe and interpret irreducible bipolar tensions constitutive of dialectical constructions.

When I tried to describe multi-dimensional and generative qualities of dialectical relationships with metaphors, I was merely one step removed from the constraints of ordinary and academic language; and even that step was an illusion. Diagrams promised fewer limitations than did written text. I wanted to create three-dimensional diagrams on disc to accompany this thesis; diagrams that would represent dialectical movement through changes in colour and position of key components.

Implications For Practice And Policy

Anne wrote the following recommendations regarding practice.
1. Teachers need time and opportunities to share with colleagues and learn how other teachers make sense of assessment.
2. Students should have more input into the assessment process.
3. Communication among parents, students, and teachers needs to be increased so that assessment practices are understood by all concerned.
4. Assessment policy should be developed from research of assessment practices in classrooms that support and promote learning.

I concur with Anne that teachers need opportunities to gain insights into how other teachers make sense of "assessment," or assessment-evaluation.

Detailed studies of how teachers make sense of assessment-evaluation is absent from the literature.

In abundance are studies of how teachers use assessment and evaluation practices, assessment and evaluation policy regarding what teachers should do, and commercially-prepared how-to-do assessment and evaluation booklets. This thesis documents one teacher's struggle to make sense of policy discussions regarding assessment and evaluation in her classroom practice. Teachers, administration, parents, and pre-service education students will gain important insights from Anne's experiences.

I also agree with Anne that policy and classroom practices regarding assessment-evaluation should be
developed in consultation with students. Children in this study had strong opinions and concerns regarding the impact of assessment-evaluation on their lives. Valuable insights are forfeited when policy-makers, researchers, parents, and teachers overlook children's standpoints. That which can be learned from children, cannot be learned from adults. A system of assessment-evaluation devoid of student input, looks like a ball game without a ball.

Much can be learned from Anne’s classroom practice. Her experiences present several important and interrelated implications for practitioners in the field. Making sense of the terms, assessment and evaluation, within the context of educational policy and classroom practices, requires sufficient time for intra- and inter-personal activity among individuals and groups of teachers. Teachers need time to work individually and collaboratively to reflect upon and discuss ideas, procedures, and activities.

Professional development days might be more appropriate times for workshops and meetings than before and after school. Many teachers are tired at the end of a school day. Some are involved in a myriad
of after-school activities like extra-curricular supervision of students, meetings with parents, staff meetings, school-based committee meetings, school board workshops, creating displays of pupil work, clean-up and organization of instructional materials, short-term and long-range planning, marking, and at some point in each term, writing report cards. Before school, some teachers are busy setting up materials, supervising extra-curricular activities, giving extra help to students, attending committee meetings, marking, and speaking with parents.

Anne's experiences indicate that making sense of assessment-evaluation involves context-specific activity. Teachers need opportunities to individually and collaboratively explore their understanding of policy and classroom practices associated with assessment-evaluation within and across contexts of their work; that is, to work collaboratively in communities of practice. The "Teacher-As-Researcher" group was one example of such a community. Within this context, Anne worked collaboratively with a group of intermediate teachers to make sense of criterion-referenced evaluation in the interests of students.
Ministry, District Office, and school administration have important roles to play in supporting communities of practice among teachers. Teachers need sufficient time and resources to work collaboratively in workshops, seminars, and conferences regarding assessment-evaluation. They also need access to print materials, technology, interdisciplinary specialists, and funds. Funding for teachers-on-call would release teachers to work collaboratively within and across classroom settings. Teachers' interests and concerns regarding assessment-evaluation need to be represented at district, provincial, national, and international conferences.

Communication within and across interest groups would be facilitated by their representation on provincial, district, and local reviews of policy and classroom practices associated with criterion-referenced assessment-evaluation. Review committees should involve Ministry personnel and a cross-section of teachers, students, parents, and principals from elementary and secondary schools. Committees should provide opportunities for new members to participate on a regular basis.
Some might conclude that an approach to criterion-referenced assessment-evaluation as a socially situated practice is costly. Maybe, maybe not. It is impossible for me to speculate what the cost might be to fund such an approach within a school district. However, I will say that the time has come for policy-makers, School Board and school administration, parents, and the public in general, to recognize in practice that the classroom teacher should be the most powerful person in the field of education. When a teacher closes the classroom door, students are at her or his mercy. In what other setting can an adult be alone with unrelated children in the same room for months at a time? The power of teachers to influence children is daunting. This power is worth investing in.

Widespread and inclusive communities of teachers working collaboratively promise new ways of thinking about criterion-referenced assessment-evaluation. In addition, teachers' communities would involve spiralling memberships of interest groups in terms of students, parents, school administrators, School Board personnel, and policy-makers. Debate would be more
inclusive, and therefore, more equitable. This approach promises consensus-building fraught with contradiction and brimming with relevancy.

I concur with Anne's recommendation that policy regarding assessment-evaluation should be informed by research on classroom practices that support and promote learning. Bachor and Anderson's (1993) study of elementary classroom assessment practices, and Bartley, Carlisle, and Erickson's (1993) work regarding performance assessment are good examples of this type of research. I also support Anne's proposal that policy-makers adopt an approach wherein policy-development is community-based and built by teachers, students, parents, administration, community members, and consolidated by the Ministry of Education.

Some might wonder how provincial government policy would account for multiple and even conflicting interpretations of assessment-evaluation. This is not a new problem. Issues of diversity are addressed regularly in debate regarding multiculturalism. An approach to criterion-referenced assessment-evaluation as socially situated practice assumes that the relevance of education policy must be understood from
the standpoints of people directly affected, and that education policy must be accountable to these standpoints.

The language of policy documents represents a serious area of concern. Language is an important mediational means in developing understanding, and/or generating confusion. The language of policy documents should be coherent and accessible to teachers, parents, students, and administration. Because these documents are for teachers, parents, and administration, they should use language of these people. Ordinary language should replace clinical and technical language.

For example, the terms, norm-referenced evaluation and criterion-referenced evaluation, were not a part of Anne's vocabulary, the vocabulary of her colleagues, and the vocabulary of parents. One way to alleviate the problem might be to dispense with reductionist and decontextualized language. Norm-referenced evaluation, criterion-referenced evaluation, and criterion-referenced letter grades, are meaningless expressions for many teachers because they are not contextually sensitive to their classroom practices.

The language of teaching is context-bound. In
contrast, policy documents used the terms, assessment and evaluation, as words with meaning beyond classroom contexts; that is, as headless horsemen. In addition, the purported distinction between assessment and evaluation was mystical to say the least. Terms like assessment and evaluation might be useful within the field of education if they are understood from teachers' standpoints.

The role of media in shaping public opinion regarding education policy poses a serious concern. The Vancouver Sun newspaper reported that parents were unhappy with the "Year 2000's" standards of achievement, and that parents wanted a return to letter grades. This statement was misleading from the point of view that parents in this study were not polled by the government for their opinions.

As long as education remains a government franchise, the power of media to influence public opinion regarding education represents the power to pull votes. Policy-makers, administration, teachers, students, and parents need to collaboratively examine the role of media in educational decision-making.
Implications For Further Research

As a result of this study, there are several suggestions that researchers may use in conceptualizing further research studies.

This case study described and interpreted policy and classroom practices associated with criterion-referenced assessment-evaluation from the standpoints of a school principal, a classroom teacher, students, and parents directly associated with one grade 5/6 class. This is valuable as a starting point to understanding assessment-evaluation as socially situated but further research should examine how it is situated within, and across, multiple-case and collaborative research projects.

Multiple studies of the situatedness of assessment-evaluation will yield important insights into peoples' interpretations of, and responses to, education policy and classroom assessment-evaluation practices within, and across, various contexts. A more inclusive and complex understanding of assessment-evaluation will result from studies of differences and similarities among peoples' interpretations and responses.
Further studies need to account for variations in gender, socioeconomic status, race, ethnicity, handicapping conditions, and geographical location. For example, researchers need to examine the relevance of education policy and classroom practices associated with assessment-evaluation within the context of socioeconomically disadvantaged communities. Socioeconomic status just might matter to people's interpretations of, and responses to, education policy and classroom practices associated with assessment-evaluation. Further research needs to account for the relationship between socioeconomic differences and assessment-evaluation.

Race, ethnicity, gender, handicapping conditions, and geographical location might also matter to peoples' understanding of, and responses to assessment-evaluation. For example, what matters to city-dwellers might have little or no relevance for people living in urban and remote communities. An approach to assessment-evaluation as situated seeks to understand the relevance of assessment-evaluation from the standpoints of diverse constituencies.

Finally, research should make an active
contribution to policy decisions. Each researcher must decide how to make her or his contribution. My objective is to share recommendations from this study regarding policy and practice with the British Columbia Ministry of Education.

In many cases assessment—evaluation has been a form of abuse and an exercise of power rather than a way to increase learning. Anne's purpose is learning. Policy-makers and parents want something else. Students are caught in the middle. Anne's purpose should take more priority.

Anne tries to promote and support students learning how to learn through her classroom assessment—evaluation procedures and practices. She encourages children to take responsibility for, and be sensitive critics of, their learning, to approach problems creatively and critically, to work collaboratively, to be honest, respectful, and considerate towards teachers, office staff, janitors, peers, parents, and community members, to volunteer time and resources to help others, and to take pride in their work. Anne does not have the power to focus solely on these qualities of character and skills in her classroom
assessment-evaluation practices. I can only imagine how a generation of students raised with Anne's ethics might change the world.
References


Baron, Mark A. & Boschee, Floyd (1995). Authentic
assessment: The Key to unlocking student success.

PA: Technomic Pub.


British Columbia, Ministry of Education (1989a). *Year
290


Capturing teachers' knowledge: Performance assessment a) and post-structuralist epistemology b) from post-structuralist perspective c) and post-structuralism d) none of the above. Educational Researcher, 23(5), 11-18.


Harding, Sandra (1993). Rethinking standpoint epistemology: What is 'strong objectivity'? In Alcoff, Linda and Potter, Elizabeth (Eds.), Feminist epistemologies (pp. 49-82). London:


Lidz, Carol Schneider (Ed.). (1987). Dynamic


Douglas K. (Eds.) What is intelligence?: Contemporary viewpoints on its nature and definitions (pp.155—162). New Jersey: Ablex Publishing.


of mind (pp. 1-34). New York: Cambridge University Press.

Appendix A

Student Research Criteria
RESEARCH PROJECT

Name: ___________________________    Date: ______________

PURPOSE: Creating plans to organize and guide the research.

CRITERIA:

For experienced student researchers:

* Plans must be available for review and reasons given why the plans are useful.

* Evaluation criteria must also be presented and explained in terms of their value to the research.

* Plans must be done in pencil.

* Evidence of the use of plans in creating the research report and development of products must be evident.

For remaining students:

* Stage 1 and Stage 2 of the planning sheets are available for review and are written in pencil.

* The Stage 1 planning sheet outlines at least four different sub-topics for the main topic of research. There is at least one question and a category written for each sub-topic.

* Information about each sub-topic is written in point form on the Stage 2 planning sheet.

* Evidence of the use of plans in creating the research report and development of products must be evident.

Self Assessment  4  3  2  1

My project deserves this mark because ____________________________

Teacher Assessment  4  3  2  1

Comment: ____________________________________________
PURPOSE: Writing to convey information in an organized manner.

CRITERIA:

* Neat, interesting Title Page
* The report addresses each sub-topic.
* The report conveys information that is written in the plans.
* The report is neatly written or typed on the computer.
* Sentence structure, use of punctuation, and spelling is correct in the report.
* Report must include reference list.

Self Assessment 4 3 2 1

My project deserves this mark because

Teacher Assessment 4 3 2 1

Comment: 
RESEARCH PROJECT

Name: ___________________________  Date: _________________

PURPOSE: Creating a product to convey information.

CRITERIA:  Product #1

* The product conveys information that is written in the plans.

* The product is neatly constructed.

* The product exhibits complexity.

* The product conveys information effectively.

* 

* 

Self Assessment  4 3 2 1

My project deserves this mark because __________________________

Teacher Assessment  4 3 2 1

Comment: ____________________________________________

_________________________________________________________________

_________________________________________________________________
RESEARCH PROJECT

Name: ___________________________ Date: ________________

PURPOSE: Creating a product to convey information.

CRITERIA: Product #2.

* The product conveys information that is written in the plans.

* The product is neatly constructed.

* The product exhibits complexity.

* The product conveys information effectively.

*

*

Self Assessment 4 3 2 1

My project deserves this mark because ____________________________

Teacher Assessment 4 3 2 1

Comment: ____________________________

_________________________________________________________________________

_________________________________________________________________________
RESEARCH PROJECT

Name: ___________________________ Date: ________________

PURPOSE: Giving an oral presentation to convey information about the research.

CRITERIA:

* Clear, projecting voice
* Eye contact
* Mostly memorized - cue cards
* Good posture
* Good pacing
* Please be polite
* 
* 

Self Assessment 4 3 2 1
My presentation deserves this mark because ____________________________________________

Teacher Assessment 4 3 2 1
Comment: ____________________________________________________________
Appendix B

Criterion-Referenced Worksheet
4. powerful/highly effective
3. capable, confident
2. partial, developing
1. limited, undeveloped

Unit performance(s) or demonstration(s)

Using a video as a new stimulus for writing to reinforce skills and build new ones.

CRITERIA

- must follow story line of video.
- must imagine the gaps in the story and write with creativity and description.
- must write from one character’s point of view and show the character’s personality.
- must have a beginning, middle and end.
- must use new paragraphs for new ideas.
- must use correct mechanics of writing.

Self Assessment 4 3 2 1

My mark is a 3 because I didn’t have one of the criteria, which was the mechanics of writing.

Teacher Assessment 4 3 2 1 Comment

Your story was interesting but you do need to work hard on the mechanics of writing.
Comments for Document 3.1.95 / March 31, 1995.

Unit performance(s) or demonstration(s):

"Using a video as a new stimulus for writing to reinforce skills and build new ones."

CRITERIA:

- must follow story line of video.
- must imagine the gaps in the story and write with creativity and description.
- must write from one character's point of view and show the character's personality.
- must have a beginning, middle, and end.
- must use new paragraphs for new ideas.
- must use correct mechanics of writing.

Self-Assessment:

My mark is a 3 because I didn't have one of the criteria which was the mechanics of writing.

Teacher-Assessment:

Comment: Your story was interesting but you do need to work hard on the mechanics of writing.
Appendix C

Classroom Performance Criteria Charts
<table>
<thead>
<tr>
<th>POWERFUL</th>
<th>CAPABLE</th>
<th>DEVELOPING</th>
<th>UNDEVELOPED</th>
</tr>
</thead>
<tbody>
<tr>
<td>• 100% effort</td>
<td>• good consistent student</td>
<td>• completed some of the work assigned</td>
<td>• work unfinished</td>
</tr>
<tr>
<td>• colorful fluid writing</td>
<td>• asks appropriate questions</td>
<td>• somewhat satisfactory overall, but parts may need work</td>
<td>• spelling incorrect</td>
</tr>
<tr>
<td>• errorfree work</td>
<td>• attractive appearance or format</td>
<td></td>
<td>• no evidence of studying</td>
</tr>
<tr>
<td>• very sophisticated</td>
<td>• consistent NCC (neat, complete, correct)</td>
<td></td>
<td>• cannot explain work</td>
</tr>
<tr>
<td>• varied representations of learning</td>
<td>• follows criteria</td>
<td>• some evidence of planning a little</td>
<td>• work carelessly done</td>
</tr>
<tr>
<td>• can teach others accurately</td>
<td>• interesting information</td>
<td>• some but not highly detailed work</td>
<td>• missing significant information</td>
</tr>
<tr>
<td>• uses highly descriptive words</td>
<td>• helpful to others</td>
<td>• able to do work but not perfectly</td>
<td>• very low effort</td>
</tr>
<tr>
<td>• covers all criteria</td>
<td>• thoughtful</td>
<td>• might have trouble explaining work</td>
<td>• did not study</td>
</tr>
<tr>
<td>• impressive finished product</td>
<td>• follows class rules when teacher isn't looking</td>
<td></td>
<td>• thoughtless, inappropriate or impolite group work</td>
</tr>
<tr>
<td>• goes above and beyond requirements</td>
<td>• good work habits</td>
<td>• tries to understand</td>
<td>• directions not followed</td>
</tr>
<tr>
<td>• varied representations of learning</td>
<td>• appealing form of presentation</td>
<td></td>
<td>• not completed on time</td>
</tr>
<tr>
<td>• outstanding format</td>
<td>• some unique aspects</td>
<td>• trying to stay on task</td>
<td>• basic criterion not met</td>
</tr>
<tr>
<td>• creative and unique in an appropriate way</td>
<td>• very high effort</td>
<td>• trying to develop new skills</td>
<td>• work difficult to understand</td>
</tr>
<tr>
<td>• catches attention of others</td>
<td>• well organized and edited</td>
<td>• trying to spend time wisely</td>
<td>• time not used wisely</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• usually listens to teacher</td>
<td>• distracts group</td>
</tr>
</tbody>
</table>

Sample Poster Phrases From Our Classroom

- many errors
- sloppy
- poorly detailed
- disorganized
- needs improvement in effort, skills or knowledge
Appendix D

Report Card
## Structured Written Report

1) Structured Written Report: 

2) Letter Grades (Explanation on reverse)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Arts</td>
<td></td>
</tr>
<tr>
<td>Mathematics</td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td></td>
</tr>
<tr>
<td>Social Studies</td>
<td></td>
</tr>
<tr>
<td>Art</td>
<td></td>
</tr>
<tr>
<td>Music</td>
<td></td>
</tr>
<tr>
<td>Physical Education</td>
<td></td>
</tr>
</tbody>
</table>

Other

Final Grade for Year

<table>
<thead>
<tr>
<th>Subject</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Arts</td>
<td></td>
</tr>
<tr>
<td>Mathematics</td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td></td>
</tr>
<tr>
<td>Social Studies</td>
<td></td>
</tr>
<tr>
<td>Art</td>
<td></td>
</tr>
<tr>
<td>Music</td>
<td></td>
</tr>
<tr>
<td>Physical Education</td>
<td></td>
</tr>
</tbody>
</table>

Assigned to Grade: June 1995

Teacher's Signature

School Administrator's Signature
MISSION

To enable students to reach their intellectual, social, aesthetic and physical potential in challenging and stimulating settings which reflect the worth of each individual and promote mutual respect, co-operation and social responsibility.

IN THE INTERMEDIATE YEARS, STUDENTS:

- study language arts, social studies, mathematics, science, information and computer studies, applied skills, physical education and fine arts
- develop their ability to solve problems, think critically, and make decisions
- develop social and personal skills, good work habits, confidence, sense of self-worth, and understanding of the value of physical and emotional well-being

Parent Involvement

Parents have a crucial role to play in the education system and their child(ren)'s progress. Regular involvement and communication with your child's school are encouraged. This can be in many ways, including parent-teacher conferences, parent meetings and school consultative committees.

This Report Includes:

1) Structured Written Report:
   • what the student is able to do
   • areas requiring further attention or development
   • ways of supporting the student’s learning
   • comments on student behaviour, attitude, work habits and effort

2) Letter Grades
   • indicate student’s progress in the curriculum
   A The student demonstrates excellent or outstanding performance in relation to the expected learning outcomes for the course or subject and grade.
   B The student demonstrates very good performance in relation to the expected learning outcomes for the course or subject and grade.
   C+ The student demonstrates good performance in relation to the expected learning outcomes for the course or subject and grade.
   C The student demonstrates satisfactory performance in relation to the expected learning outcomes for the course or subject and grade.
   C- The student demonstrates minimally acceptable performance in relation to the expected learning outcomes for the course or subject and grade.
   D The student demonstrates unsatisfactory performance in relation to the expected learning outcomes for the course or subject and grade.
   E Probable failure.

Letter Grades for Final Report
A, B, C+, C, P = Pass, SG = Standing Granted, F = Fail

Parents: Please keep this copy and return the report card envelope only. Thank You.
Appendix E

Martin's Flower
4. powerful/highly effective
3. capable, confident
2. partial, developing
1. limited, undeveloped

CRITERIA

Self Assessment  4 3 2 1

My _____________________________ is _____________________________
because _______________________________________________________

Teacher Assessment  4 3 2 1  Comment _____________________________
Appendix F

Report Card Reflections
Report Card Reflections

In general, how did you feel about your marks and comments?

____________________________________

____________________________________

____________________________________

Language Arts

I feel good about________________________________________

____________________________________

____________________________________

I need to improve on________________________________________

____________________________________

____________________________________

My plan is to________________________________________
**Math**

I feel good about


I need to improve on


My plan is to


**Science**

I felt good about


I need to improve on


My plan to


Social Studies

I felt good about

I need to improve on

My plan is to

Art

I felt good about

I need to improve on
My plan is to

I felt good about

I need to improve on

My plan is to

French

I felt good about

I need to improve on

My plan is to
Appendix G

Glossary
Assessment - 1. "The systematic process of gathering information about student learning, what they know, are able to do, and are trying to do" (Victoria, February 1994, p.23); 2. Collecting information about student learning and judging how effectively students demonstrate their understanding of what they are learning, or have learned, according to the curriculum (Anne).

Assessment-Evaluation - Ongoing and interconnected classroom procedures and practices for gathering information and making judgements about student progress based upon curriculum criteria and expected learning outcomes (Author).

Authentic Assessment - An approach to assessment wherein practices are designed to evaluate student performance within the context of real-life tasks (Baron & Boschee, 1995).

Capital - Power (Bourdieu, 1986a).

Criterion-Referenced Evaluation - 1. An approach to evaluation that compares student performance with curriculum criteria (Victoria, September 1994); 2. An instructional approach that involves student participation in setting and defining performance criteria (Anne, 1995).


Dynamic Assessment - An approach to assessment that focuses on what students are capable of learning (Lidz, 1987).

Economic Capital - Power in the form of financial resources (Bourdieu, 1986a).

Evaluation - "The process of making judgements and
decisions based on the interpretation of evidence of student learning gathered throughout assessment. Evaluation might be done by the teacher or the student independently or in collaboration" (Victoria, February 1994, p.23).

Field - Arena of unequal relations of power among people competing for capital (Bourdieu, 1977).

Genetic Structuralism - Pierre Bourdieu's (1990) dialectical interpretation of agency and structure.

Habitus - Mediating construct bridging agency and structure. Individual and group habitus involve dispositions that guide production, reproduction, and structuration of practices within and across fields (Bourdieu, 1977).

Mediated Action - "Action interpreted as involving an irreducible tension between mediational means and the individuals employing these means" (Wertsch, 1991, p.64).

Mediated Agency - Individuals using mediational means in sociocultural settings (Wertsch & Rupert, 1993).

Mediational Means - Culturally, historically, and institutionally formed languages and tools that bridge people's actions and sociocultural settings (Wertsch, 1991).

Norm-Referenced Evaluation - "An approach that compares student performance with the performance of other students" (Victoria, September 1994, p.23).

Performance Assessment - "Assessment that is based on students demonstrating what they can do" (Victoria, February 1994, p.23).

Portfolio - "A purposeful collection of student work that shows the student's effort, progress, or achievement over time" (Victoria, February 1994, p.23)
Portfolio Assessment - "A method of gathering student work samples for the purpose of evaluating student knowledge and learning" (Victoria, February 1994, p.23).

Practice - 1. Dialectical production of habitus and field (Bourdieu, 1990); 2. Neither Anne nor her colleagues used the term.

Reporting - The presentation of information gained through evaluation based on identified standards (Victoria, February 1994).

Self-Assessment / Self-Evaluation - "The ongoing process in which students get to know themselves as learners by reflecting on their own performance, products, thinking, and learning" (Victoria, February 1994, p.23).

Social Capital - Power in the form of membership in a group or organization wherein social connections have high exchange value for other types of capital (Bourdieu, 1986a).

Sociocultural Theory - This thesis uses Jim Wertsch's (1991) interpretation of sociocultural theory wherein the relationship between human action and sociocultural setting is interpreted as dialectical.

Student-Centred Conference - "A meeting of teacher, parent(s) and student to discuss the student's learning" (Victoria, February 1994, p.23).

Symbolic Capital - Power to generate recognition from misrecognition (Bourdieu, 1990).

Zone of Proximal Development - The interval between what an individual can do independently in a problem-solving situation and what the individual can do in a problem-solving situation with assistance from an adult or by collaborating with more adept peers (Vygotsky, 1978; 1986).