



VCU

Virginia Commonwealth University
VCU Scholars Compass

MERC Publications

MERC (Metropolitan Educational Research
Consortium)

1999

Teachers' Classroom Assessment and Grading Practices: Phase 2

James H. McMillan

Virginia Commonwealth University, jhmcmill@vcu.edu

Daryl Workman

Virginia Commonwealth University

Follow this and additional works at: http://scholarscompass.vcu.edu/merc_pubs



Part of the [Education Commons](#)

Downloaded from

http://scholarscompass.vcu.edu/merc_pubs/57

This Research Report is brought to you for free and open access by the MERC (Metropolitan Educational Research Consortium) at VCU Scholars Compass. It has been accepted for inclusion in MERC Publications by an authorized administrator of VCU Scholars Compass. For more information, please contact libcompass@vcu.edu.

TEACHERS' CLASSROOM ASSESSMENT AND GRADING PRACTICES:

Phase 2

James H. McMillan, Professor
Virginia Commonwealth University

Daryl Workman, MERC Research Fellow
Virginia Commonwealth University

Copyright©1999. Metropolitan Educational Research Consortium (MERC), Virginia Commonwealth University

*The views expressed in MERC publications are those of individual authors and not necessarily those of the Consortium or its members.

Executive Summary

Classroom assessment and grading practices are becoming a greater focus of educational inquiry as teachers and policymakers become more accountable to the public for educational outcomes via assessment driven instructional practices. This study is an attempt to better understand the classroom assessment and grading practices of teachers, which have previously been described as a hodgepodge mix of student attitude, effort and achievement. Specifically, the following questions regarding teachers' assessment and grading practices are addressed:

- What is the current state of assessment practice and grading by teachers?
- What assessment and grading topics do teachers identify as needs to be addressed in in-service?
- What is the relationship between assessment and grading practices and grades given to students?
- What are the relationships between grade level, and subject taught and assessment and grading practices?
- What are the reasons teachers give for their assessment and grading decision-making?
- What is the impact of the SOL tests on the extent to which different assessment techniques are used in the classroom?
- What classroom assessment and in-service needs do teachers have?

Results of the research indicates that teachers do in fact use a multitude of factors to assess and grade students, including academic performance, effort, improvement, ability, homework, and extra credit. However, this study attempts to look beyond a "hodgepodge" explanation of assessment and grading practices in order to uncover relationships that help to further explain teachers' assessment and grading practices and decision-making processes.

This report summarizes the findings from Phase 2 of the study, which focused on personal interviews with teachers. Phase 1 of this research, which was reported in Fall, 1998, surveyed 921 elementary, 597 middle and 850 high school teachers. A total of 28 mostly middle and high school mathematics and English teachers were interviewed individually.

The analysis of Phase 2 data indicates that there is tension between two sources of influence on teacher decision-making concerning assessment and grading practices. One source is teacher beliefs and values while a second source is external pressures and constraints. Teacher beliefs and values focus on assessment and grading practices that will encourage and support student learning. Teachers "pull" for students, devising approaches to assessment and grading that make it likely that students will succeed. Assessment and grading practices tend to be individualized to a certain extent for different students, and used as a way to keep students motivated and engaged. Teachers want students to understand and learn, and want assessments that help this outcome. Constructed-response assessments are seen as providing the best information to help students succeed.

Outside pressures and constraints include parental demands and informing parents of student progress, division policies, skills needed by students once they graduate, practical constraints such as having over 100 students, and perhaps most importantly, state mandated high stakes multiple choice testing. It appears that the state testing program has become a significant influence on teacher decision making, lessening to some extent assessment and grading practices that more clearly, from the teachers' perspectives, promote student learning.

Implications for teacher professional training and development are made in light of the tension between these two sources influences.



Table of Contents

Preface	1
Introduction	2
Review of Literature	4
Research Questions	13
Methodology	14
Findings	20
Conclusions	44
Implications	46
References	49
Appendix A	52

Preface

The research in this report was directed by a team of individuals. This team identified the research problem and questions, developed a research design, assisted in gathering data from teachers, and took an active role in identifying samples and analyzing data. The principal investigators are grateful for their contribution and assistance. The members of the research team include the following:

James McMillan
Virginia Commonwealth University

Catherine Nolte
Henrico County Public Schools

Daryl Workman
Virginia Commonwealth University

Yvonne Smith-Jones
Hopewell City Public School

Lin Corbin-Howerton
Chesterfield County Public Schools

Stephanie Couch
Pocahontas Middle School
Powhatan County Public Schools

Joseph Tylus
Monacan High School
Chesterfield County Public Schools

Charmaine Brooks
Westover Hills Elementary
Richmond City Public Schools

Ann Williams
Colonial Heights Middle School
Colonial Heights Public Schools

Pat Janes
Armstrong High School
Richmond City Public Schools

James Bagby
Hanover County Public Schools

Sue Jones
Patrick Copeland Elementary
Richmond City Public Schools

Carole Urbansok-Eads
Hanover County Public Schools

Audrey Johnson
Thomas Jefferson High School
Richmond City Public Schools

Deborah Pittman
Cold Harbour Elementary School
Hanover County Public Schools

Richard Williams
Richmond City Public Schools

Introduction

A significant amount of recent literature has focused on classroom assessment and grading as essential aspects of effective teaching. There is an increased scrutiny of assessment as indicated by the popularity of performance assessment and portfolios, newly established national assessment competencies for teachers (Standards, 1990), and the interplay between learning, motivation, and assessment (Brookhart, 1993, 1994; Tittle, 1994). In Virginia, the Standards of Learning and associated tests highlight the importance of assessment.

Previous research documents that teachers tend to award a “hodgepodge grade of attitude, effort, and achievement” (Brookhart, 1991, p. 36). It is also clear that teachers use a variety of assessment techniques, even if established measurement principles are often violated (Cross & Frary, 1996; Frary, Cross, & Weber, 1993; Plake & Impara, 1993; and Stiggins & Conklin, 1992).

Given the variety of assessment and grading practices in the field, the increasing importance of assessment, the critical role each classroom teacher plays in determining assessments and grades, and the trend toward greater accountability of teachers with state assessment approaches that are inconsistent with much of the current literature, there is a need to (1) understand current assessment and grading practices, (2) understand the relationship of these practices to grades given by teachers, (3) determine if “standards” teachers use to assign grades differ from one classroom to another and one school to another, (4) examine the consequential validity of the

new SOL tests on classroom assessment practices, and (5) determine assessment and grading topics that, according to teachers, need in-service.

The fourth need is related to a recently expanded conception of test validity that includes what has been called “consequential validity” or “consequential bias” (Messick, 1989; Moss, 1992). Essentially, test developers and users need to be sensitive to how assessments influence instructional practices and curriculum. The importance of consequential validity is indicated by its inclusion in the new Standards for Educational and Psychological Testing. Of interest in the current study is the effect the new statewide assessment program may have on instructional practices. For example, the assessments may result in teachers stressing a particular method of instruction or classroom testing that is consistent with the emphasis and approach adopted in the statewide system

There is a need to provide information that addresses issues of consistency and fairness in assessment and grading across classrooms and schools, to illustrate to teachers the nature of current practice and provide a stimulus for discussion, and to establish assessment and grading policy. There is also a need to understand the motivation and reasons for using specific assessment and grading practices.

The purpose of both Phases of this investigation, then, was to describe the classroom assessment and grading practices of teachers, determine if meaningful relationships exist between these practices and grade level, subject matter, ability levels of different classes, and to understand the reasons teachers give for using certain assessment and grading practices, and to document teacher needs for inservice education related to assessment. This report focuses on Phase 2 of the research.

Review of Literature

Despite the growing importance of classroom assessment and the introduction of new methods of assessment, there is relatively little research on the nature and effects of classroom assessments on student learning and motivation (Stiggins, 1997). Most assessment research has focused on standardized testing, despite evidence that teachers spend considerable time assessing students, and that student well-being is influenced by the quality of assessments given by the teacher (Stiggins and Conklin, 1992). Also, there is little empirical research on classroom assessments, with measurement experts tending instead to pay much more attention to large scale testing than classroom assessment. It is also evident that many teachers lack assessment competency (Plake and Impara, 1997). This isn't too surprising, however, since less than 50% of the teacher certification programs in the United States require no measurement course (Schafer, 1993). This remains the case, despite the fact that teacher standards for assessment competency were identified in 1990 (AFT, NCME, NEA, 1990).

Prior to the mid 1980s the literature on educational assessment focused almost exclusively on large-scale standardized testing. According to Stiggins and Conklin (1992), most inquiry on classroom assessment was based on a conceptualization similar to what had been developed for standardized testing, emphasizing paper and pencil, multiple choice testing. Furthermore, the only written standards for assessment, *Standards for Educational and Psychological Testing*, dealt primarily with standardized tests. Finally, during the 1980s the emerging literature about teacher decision-making, teacher behavior, and student achievement found little on how

classroom assessments relate to teaching or learning. Shulman (1980) concluded that most of the paper and pencil tests used for assessment were inconsistent with, and often irrelevant to, the realities of teaching. Haertel, et al. (1984), in a review of research on high school testing, concluded that little is known about teachers' or students' perceptions of the impacts of classroom assessment.

Phye (1997) states that "it is not only the assessment option that determines what we get as evidence of learning or achievement. How we use the assessment instruments or techniques also determine the nature of the knowledge a student is demonstrating. *How* we assess determines *what* we get" and thus classroom learning and assessment "go hand in hand" (p.51).

Airasian (1984) reviews literature that suggests teachers focus their classroom assessments in two areas: academic achievement and social behavior. The importance of these factors varies with grade level, with elementary teachers placing greater importance on social behavior. Airasian also found that teachers' informal "sizing up" assessments remain relatively stable throughout the year and influence student self-perceptions of ability.

Fleming and Chambers (1983), in a study that analyzed nearly 400 teacher-developed classroom tests, came to several conclusions:

- Short-answer questions are used most frequently.
- Essay questions are avoided, representing slightly more than 1% of test items.
- Matching items are used more than multiple choice or true false items.
- Most test questions, approximately 80%, sample knowledge of terms, facts, and rules and principles (94% for middle school teachers, 69% for high school teachers, and 69% of elementary school teachers).
- Few test items measure student ability to apply what they have learned.

Research by Carter (1984), in which the test development skills of high school teachers were studied, in support of what Fleming and Chambers found, reported that the teachers had

considerable difficulty recognizing or writing items that tapped "higher order" thinking skills, such as application. Stiggins and Conklin (1992), with a sample of thirty-six teachers, found that recall knowledge items were used approximately fifty percent of the time.

There is ample evidence to suggest that many teachers do not have sufficient knowledge and skill to develop, apply, and summarize classroom assessments. In a survey of 228 teachers from four grades (2, 5, 8, and 11), Stiggins and Conklin (1992) report that nearly three fourths of the teachers indicated some concern about their own tests. Examples of the kinds of concerns expressed included: "Are my tests effective? How can I make them better? Do they focus on students' real skills? Are they challenging enough? Do they aid in learning?" (p. 39). Concern was greatest for high school teachers. Only 15% of high school teachers indicated that they had no concerns about their assessments. Stiggins and Conklin also asked 24 teachers to keep a journal to reflect upon their assessment practices. The analysis focused on how teachers describe their assessments and what specific issues were raised related to their assessments. They found that teachers were most interested in assessing student mastery or achievement, and that performance assessment was used frequently. The nature of the assessments used in each class was coupled closely with the roles each teacher set for her students, teacher expectations, and the type of teacher-student interactions desired. The results of these investigations led to the development of classroom assessment profiles. The profile was tested with eight high school classrooms, resulting in the following key factors:

- Assessment purposes
- Assessment methods
- Criteria used in selecting assessment methods
- Quality of assessments
- Feedback to students
- Teacher as assessor (background, preparation)

- Teacher perception of the students
- The assessment-policy environment

These components can be used to characterize diverse assessment practices and environments.

Two recent studies document teacher beliefs and knowledge about classroom assessment.

Frary, Cross, and Weber (1993) used a statewide random sample of 536 high school teachers of academic subjects to survey self-report practices and beliefs about classroom assessment.

Frequency of use of various kinds of test questions revealed the following percentages:

<u>Type of Question</u>	<u>Seldom or Never</u>	<u>Frequently or Always</u>
Short answer	17%	56%
Essay	41%	38%
Multiple choice	21%	52%
True-false	47%	19%
Performance	30%	37%

These results suggest that teachers use a variety of assessment approaches. The teachers were asked to indicate degree of agreement to many statements concerning grading and assessment practices. Concerning assessment, it was noteworthy that 66% of the teachers agreed that essay tests provide a better assessment of student knowledge than do multiple choice tests; that 47% agreed that the nature of multiple choice items encourages superficial learning; and that better measurement occurs when teachers award partial credit rather than scoring simply right or wrong.

A second survey of teachers, taken in 1992, was structured to obtain teacher competency concerning assessment practices by asking teachers to indicate which of several possible answers to assessment questions was best (Plake and Impara, 1997). A national random sample of 555 elementary, middle, and high school teachers was used. Overall mean performance on the survey

was 66% correct. Teachers did better on items related to choosing and administering assessments and significantly worse on communicating results. According to the authors, the results "give empirical evidence of the anticipated woefully low levels of assessment competency for teachers" (p.67). The results also showed that teachers who had had a measurement course performed better than teachers who lacked this background.

In summary, the small amount of existing literature on classroom assessment practices indicates that teachers probably need further training to improve the quality of the assessments that are used. There continues to be reliance on selected-response tests, with conflicting evidence concerning the use of essays. Whatever the type of question, few are written to tap students' higher level thinking skills. Appropriately, teachers appear to use a variety of assessment methods. There is clearly a need for more research on classroom assessments. Classroom assessments consume significant amounts of time for both teachers and students, and have important consequences. Particularly absent in the literature are examination of relationships between classroom assessment practices and grading, how teachers use assessments to set standards, and how teachers make decisions about the assessments they use.

Teachers' grading practices have received far more attention in the literature than have assessment practices. This may be due to the salient and summative nature of grades to students and parents. Grades have important consequences and communicate student progress to parents.

A study by Stiggins, Frisbie, and Griswold (1989) set the stage for research on grading by providing an analysis of current grading practices as related to recommendations of measurement specialists and newly established Standards for Teacher Competence in Educational Assessment of Students (American Federation of Teachers, National Council on Measurement in Education, National Education Association, 1990). In this study the authors interviewed and/or observed 15

teachers on 19 recommendations from the measurement literature. They found that teachers use a wide variety of approaches to grading, and that they wanted their grades to both fairly reflect student effort and achievement, as well as to motivate students. Contrary to recommended practice, it was found that teachers value student motivation and effort, and set different levels of expectation based on student ability.

Brookhart (1994) conducted a comprehensive review of literature on teachers' grading practices. Her review identified 19 studies completed since 1984. Seven studies investigated secondary school grading, 11 studies both elementary and secondary, and one study elementary teachers. Three general methods of study were identified: surveys in which teachers responded to questions concerning components included in grading, grade distributions, and attitudes toward grading issues; surveys in which teachers were asked to respond to grading scenarios, asking what they would do in various circumstances; and qualitative methods, including interviews, observation, and document analysis. Despite methodological and grade level differences, the findings from these studies are remarkably similar. This suggests that conclusions warranted from the research are generalizable. Taken together, Brookhart comes to the following conclusions:

- Teachers inform students of the components used in grading.
- Teachers try hard to be fair in grading.
- Measures of achievement, especially tests, are major contributors to grades.
- Student effort and ability are used widely as components of grades.
- Elementary teachers rely on more informal evidence and observation, while secondary teachers use paper and pencil achievement tests and other written evidence as major contributors.
- Teachers' grading practices vary considerably from one teacher to another, especially in perceived meaning and purpose of grades, and how nonachievement factors will be considered.
- Teachers' grading practices are not consistent with recommendations of measurement specialists, especially confounding effort with achievement.

In one study, Brookhart (1993) investigated the meaning teachers give to grades and the extent to which value judgments are used in assigning grades. The results indicated that low ability students who tried hard would be given a passing grade even if the numerical grade were failure, while working below ability level did not affect the numerical grade. That is, an average or above average student would get the grade earned, whereas a below average student gets a break if there is sufficient effort to justify it. Teachers were divided about how to factor in missing work. About half indicated that a zero should be given, even if that meant a failure for the semester. The remaining teachers would lower the grade but not to a failure. The teachers' written comments showed that they strived to be "fair" to students. Teachers also seemed to indicate that a grade was a form of payment to students for work completed. More comments indicated that grades were something students earned as compared to grades indicating academic achievement, as compensation for work completed. This suggests that teachers, either formally or informally, include conceptions of student effort in assigning grades. Because teachers are concerned with student motivation, self-esteem, and the social consequences of giving grades, using student achievement as the sole criteria for determining grades is rare. This is consistent with earlier work by Brookhart (1991), in which she pointed out that grading often consists of a "hodgepodge" of attitude, effort, and achievement.

Cross and Frary (1996) report similar findings concerning the "hodgepodge" nature of grades. They surveyed 310 middle and high school teachers of academic subjects in a single system. A teacher survey was used to describe grading practices and opinions regarding assessment and grading. Consistent with Brookhart, it was reported that 72% of the teachers raised the grades of low ability students. One-fourth of the teachers indicated that they raise grades for high effort

"fairly often." Almost 40% of the teachers indicated that student conduct and attitude were taken into consideration when assigning grades. Interestingly, a very high percentage of teachers agreed that effort and conduct should be reported separately from achievement. Over half of the teachers reported that class participation was rated as having a moderate or strong influence on grades.

An earlier statewide study by Frary, Cross, and Weber (1993), using the same teacher survey that was used by Cross and Frary (1996), found similar results. Percentages of teachers agreeing or tending to agree to the following statements illustrate this conclusion:

<u>Item</u>	<u>Percentage</u>
• A student's ability should be taken into consideration in awarding final grades.	66
• An exceptionally low or high degree of student effort should be recognized by adjustment of the final grade.	66
• The amount of knowledge a student <i>gains</i> over the instructional period should be taken into consideration in awarding the final grade.	85
• Laudatory or disruptive classroom behavior should be considered in determining final grades.	31
• The minimum passing score on a test should be based at least in part on the scores earned by students of marginal ability who have been putting forth satisfactory effort.	64

Another recent study by Truog and Friedman (1996), further confirms the notion of hodgepodge grading. In their study the written grading policies of 53 high school teachers were analyzed in relation to grading practices recommended by measurement specialists, and a focus group of eight teachers was conducted to probe reasoning used by the teachers. The study was based on an earlier investigation by Stiggins, Frisbie, and Griswold (1989). Friedman and Manley (1991) also found that teachers routinely use ability, attitude, effort, participation, and other factors in addition to achievement when determining grades. Truog and Friedman (1996)

found that written policies were consistent with earlier studies of teacher beliefs and practice. Nine percent of the teachers included ability as a factor in determining grades, 17% included attitude, 9% included effort, 43% included attendance, and 32% included student behavior.

Another survey of 143 elementary and secondary school teachers conducted by Cizek, Fitzgerald and Rachor (1995) collected data on teachers' assessment-related practices. Results indicated that assessment practices "were highly variable and unpredictable from characteristics such as practice setting, gender, years of experience, grade level or familiarity with assessment policies in their school district" (p. 159). Furthermore, teachers generally use a variety of objective and subjective factors to maximize the likelihood that students obtain good grades. Overall, the authors concluded that "many teachers seemed to have individual assessment policies that reflected their own individualistic values and beliefs about teaching" (p.160). The authors argue that grades should be used in more meaningful ways to communicate about student performance.

In summary, the literature on grading strongly supports the notion that teachers believe it is important to combine nonachievement factors, such as effort, ability, and conduct, with student achievement, to determine grades. While the studies are clear in this conclusion, less is known about how teachers decide to weigh these nonachievement factors in determining grades. Also, many of the surveys and other approaches in previous studies have asked teachers about their beliefs or projected behavior based on scenarios. It is possible that actual grading practice may be different. Despite increased focus on assessment and teacher competence with respect to measurement and grading, there appears to be a continuing discrepancy between recommended practice and teacher beliefs about grading. Furthermore, while descriptions of grading practices

are plentiful, there is little research on the relationship between grading practices and student motivation and achievement.

The literature reviewed on the nature and effect of assessment and grading practices on student achievement has demonstrated that there is little empirical evidence of the specific effects of using particular assessments and grading procedures. This is due in part to the complex nature of teaching, and how assessment and grading are only a part of instruction. Assessment and grading continue to be a private activity, with considerable variation among teachers. While "newer" forms of assessment, such as performance-based and portfolio, are based on recent research on cognitive learning, the suggestions are based on theory and not empirical evidence. There are several studies which show that teachers engage in assessment and grading practices that are not consistent with what would be recommended by measurement "experts." For example, combining nonachievement factors like effort, ability, and conduct with student achievement to determine grades, as well as "hodgepodge" grading. While descriptions of grading practices are plentiful, there is little research on the relationship between grading practices and student motivation and achievement. One theoretical model postulated by Brookhart (1997) represents an initial perspective about how assessment and grading practices affect self-efficacy, effort, and achievement. There is a strong research base with respect to the two major contributors to motivation (self-efficacy and importance, utility, and value), but not much about how specific assessment and grading practices effect these two components.

Research Questions

The purpose of the proposed research (both phases) is to gather information from teachers regarding their assessment and grading practices to answer the following questions:

- What is the current state of assessment practice and grading by teachers?
- What assessment and grading topics do teachers identify as needs to be addressed in in-service?
- What is the relationship between assessment and grading practices and grades given to students?
- What are the relationships between grade level, and subject taught and assessment and grading practices?
- What are the reasons teachers give for their assessment and grading decision-making?
- What is the impact of the SOL tests on the extent to which different assessment techniques are used in the classroom?
- What classroom assessment and in-service needs do teachers have?

Forshawdowed research questions guiding Phase 2 include:

1. What is the nature of teacher decision making concerning classroom assessment and grading practices?
2. What influences teacher decision making concerning classroom assessment and grading practices?
3. What classroom assessment and grading practices are identified?
4. What justification do teachers give for their classroom assessment and grading practices?

Methodology

Research Design

The research consisted of two phases, one involving a written survey of a large number of teachers and one using face to face interviews. Phase 1 included development and administration of a teacher questionnaire to survey teachers' assessment and grading practices and in-service needs. Quantitative analysis of the data included data reduction, descriptive statistical results, and the investigation of relationships with analysis of variance and correlational procedures. Phase 2 used interviews with selected teachers to investigate decision-making and justification for specific assessment and grading practices. This report is concerned with Phase 2 of the research. A qualitative research design was utilized.

Participants

On the written survey from Phase 1 teachers were given the option to participate in an in-depth interview regarding their responses to the survey. Volunteers' surveys were pulled and reviewed for maximum variation in item response. Surveys in which maximum variation responses were consistently observed were selected for further interviews. Maximum variation responses were identified via the survey Likert scale in which a response of 1 indicated "not at all" and a response of 6 indicated "extensively." Sixty (60) surveys were originally identified as meeting the maximum variation criteria for selection. The 60 teachers were then contacted by telephone and a letter was faxed to their school requesting an interview date and time. Of the 60 teachers originally selected for an interview, 28 ultimately participated in the interview process.

Of the teachers interviewed for this study, English/language arts classes were represented by teachers from more than a dozen different schools in 7 different school districts. Grades represented by the teachers were 5, 6, 7, 8 and 12, with students of varying academic abilities (low, moderate to advanced placement).

Math classes were also represented by teachers from more than a dozen different schools in 7 different school districts. Grades represented by the teachers were 7, 8, 9, 10, 11 and 12 with students of varying academic abilities (low, moderate, to advanced placement as well).

Table 1 shows the breakout of teachers according to subject matter, grade level and ability level.

Table 1
Summary of Characteristics of Teachers Interviewed

	<u>Grade Level</u>								
	<u>5</u>	<u>6</u>	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>	<u>11</u>	<u>12</u>	
Eng/LA Ability									
Low		1							
Mod	1	1	1						
High		1		2					4
Math Ability									
Low			2	2	1				
Mod				2	1	1			
High			1		1		2	2	

Interview Protocol and Process

Four members of the research team participated in the interviewing process. All but four of the interviews were tape-recorded. Interviewer notes were taken on the four unrecorded interviews, as well as for some of the tape-recorded interviews. Interviews lasted 45 to 60 minutes and took place in the teachers' schools. An interview guide was developed by the research team prior to the interviewing process. It was used in four pilot interviews and revised by the team prior to completing the sample of 28 interviews. A copy of the interview guide is attached in Appendix A.

Data Analysis

Each tape recorded interview was transcribed into approximately 8 to 14 pages of single spaced typed text. The text was then loaded into HyperRESEARCH qualitative software. Data were reviewed by team members and then coded by the university team members according to the emerging topics of the interviewees, as well as the pre-established topics identified in the interview guide. Forty nine (49) codes were initially established, although 26 became the major categories used throughout the data analysis. Following coding, members reconvened to review the coding and categorizing. The categories were then further collapsed into four themes that explain the data. Table 2 shows the initial codes identified during the data analysis, as well as their level of frequency throughout the data (codes with an asterisk [*] beside them indicate major categories).

Table 2
Summary of Codes and Frequencies of Responses

<u>Code</u>	<u>Frequency</u>
Administrative	1
Advice to other teachers	2
*Assessment rationale	53
Assessments drive lesson plans	9
*Borderline grades	18
*Choice of assessments	107
Class type	34
*District grading policy	19
*Effort versus ability	24
*Extra credit	21
*Feedback	46
*Formal versus informal assessments	11
*Grade challenged by parent	2
*Grade distributions	36
*Grades	29
*Grading policy	50
*Grading policy rationale	25
Group work	8
*Homework	4
*Influences on mode of teaching	17
*Informativeness of assessments	30
Lesson plans	2
*Lesson plans drive assessment	27
*Modification of assessments	13
*Modify assessments for spec. students	18
Objective assessments	21
*Ongoing assessments	9
*Other teacher grading policy	12
Performance assessments	1
Pre-assessments	40
Publisher made assessments	8
*Pulling for students	39
Quiz grades	6
Revision of assessments	23
Socio-economic status	1
*SOLs	52
Standardized tests	2
Student effort performance	22
*Student motivation	57
Student performance on quizzes	19
Student recall knowledge	21
Summative evaluation	1
Teacher made tests	7
Timing of assessments	7
Types of test items	2
When assessments used	2
*Worth and value of assessments	6
*Zeros	22

The selection of codes into categories was not based on frequency counts alone, but also depended upon the similarity a code had to other codes which were becoming major categories. For example, the code “grade challenged by parent” has a frequency count of only 2, yet it was still retained as a category because of its similarity to other “grade” related codes and categories.

Throughout the process of coding and categorizing, an on-going inductive analysis of the data was also occurring. The researchers were constantly interpreting the data as it unfolded in an effort to understand themes as they emerged. During this process of inductive analysis, coding and categorizing decisions were also affected such that codes and categories that initially promised explanatory value for the data were eventually discarded as new, more meaningful codes and categories developed over time. For example, the code “student recall knowledge” has a frequency count of 21, yet it did not become a major category because the on-going inductive analysis revealed, over time, that the code did not hold much explanatory value for the data.

Following the reduction of 49 codes down to 26 categories, the categories were further reduced to four (4) themes: (1.) Teacher beliefs and values, (2.) External Factors, (3.) Teacher decision-making rationale, (4.) Assessment and grading practices. These four (4) themes have powerful explanatory value for the data and offer a tentative theory of teachers’ assessment and grading practices.

As a validity check, 50% of the coded transcripts were peer reviewed by a MERC associate to determine agreement on the selection of codes assigned to chunks of data. Of 520 codes assigned to the data by the researcher, the peer reviewer agreed with 450 of them. This resulted in an 87% rate of agreement between the researcher and the peer reviewer.

Findings

The results will be presented by first explaining an overall model of teacher decision making that represents a synthesis of external factors that influence teachers and teachers' beliefs and values. Following the model, quotes from teachers that represent their beliefs and values, external factors and their decision making rationales will be presented.

A Model of Factors Influencing Teacher Assessment and Grading Practices Decision-Making

Inductive data analysis resulted in a tentative model that explains how and why teachers decide to use specific assessment and grading practices. The main tenet of the model holds that there is tension between internal, beliefs and values of teachers and external factors that are imposed on them. This tension between these two types of influences is apparent in the explanations teachers give for their assessment and grading practices. Such practices are influenced most heavily by internal beliefs and values. External pressures, especially recent SOL testing, forces teachers to use assessment and grading practices that probably would otherwise have little impact. Greater tension arises when external pressures increase, and lessens as teachers gain experience.

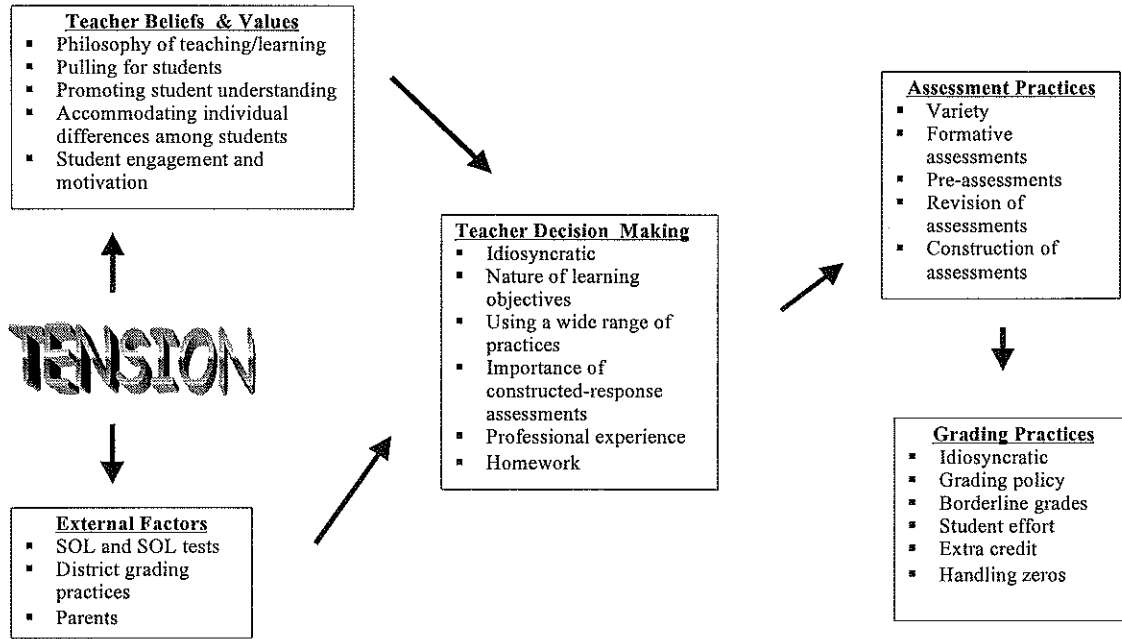
Within this model external factors are considered to be accountable (to systems, parents and students), with their end result revealing an "objective" way of documenting student performance. Teachers' beliefs and values, on the other hand, are internalized factors that are frequently idiosyncratic beliefs that comprise teachers' philosophy of education. They usually consist of beliefs that students do not fit into highly organized, objective, categories, but rather are individuals in need of flexible assessment and grading schemes. As a result, teacher

decision-making rationales are varied, individualized justifications for the types of assessments used and the grade assigned to a student.

Figure 1 illustrates the model of teacher decision making. It shows how the decision making process is influenced by teacher beliefs and external factors, leading to rationales for specific assessment practices, followed by grading practices. The model focuses on the decision making process, but both teacher beliefs and values, and external pressures, will influence assessment and grading practices in a fairly direct way. In the end, to better understand teacher decision making, it is necessary to explore factors in each of the two major categories of influences.

Figure 1. A Model of Teacher Assessment and Grading Practices Decision-Making.

A Model of Teacher Assessment & Grading Practices



Teacher Beliefs and Values

The most salient internal factor that appears to influence teacher decision-making concerning classroom assessment and grading practices is the teacher's philosophy of teaching and learning that provides justification for the practices. This internal set of values is important because it provides a rationale for using assessment and grading practices that are most consistent with what the teacher believes is most important in the teaching/learning process. For some of the teachers interviewed the philosophy of learning was focused on doing whatever was needed to help students succeed, to "pull for students." In extreme cases, this meant significant modifications of assessments, such as writing multiple forms of tests to accommodate different students' needs and abilities, allowing creative expressions such as artwork to substitute for

regular paper and pencil tests, and allowing students to veto certain types of tests questions if they feel incapable of responding to them. Other teachers were less accommodating but still indicated a philosophy based on student success. These teachers, for example, would accept late work or revisions of work.

Essentially, it's as if assessment and grading practices are whatever will best serve the purposes that are linked to a larger, more encompassing philosophy of education. For example, teachers believe that students need to be meaningfully engaged in learning, and would use assessments and grading factors that would enhance this engagement. Five categories of teacher beliefs and values were identified: Philosophy of teaching/learning, pulling for students, promoting student understanding, accommodating individual differences, and student engagement and motivation.

Philosophy of Teaching/Learning

Note how the following excerpts from the teachers frame their assessment and grading practices in a larger philosophy of teaching/learning:

- I weigh more on homework ...there are worksheets that I use. And to me, my philosophy of education is run by Dewey. The more you practice something the better, the more proficient you become in that skill.
- The daily grades are my way of making sure that they continue this process like it's supposed to be. If they haven't done the daily things that I've asked them to do then they are not going to be able to do that end result.
- I always assess early on to see what people know so that I could split groups and remediate or accelerate as needed.
- To me grades are extremely secondary to the whole process of what we do. I have goals to what I want to teach and I use assessment so that I know what I need to work on. What people have mastered and what they haven't.

Pulling For Students

For most of the teachers it was evident that they wanted very much for students to succeed and to obtain good grades. We have labeled this value “pulling for students.” It is manifest in assessment and grading practices that are designed to give students the best opportunity to be successful. In some cases it almost appeared as if teachers were using specific practices so students could pull up low test scores. Working within the constraints of the grading system, teachers wanted very much for students to do well. Here are some illustrations of this tendency:

- Maybe this is the dedicated teacher’s syndrome or whatever. I’ll chase the kid around for a long time so I can get a few points.
- I’m always trying to find some ways so that all the children can find success, not just Johnny and Suzy getting the A but also Sally and Jim can get an A also.
- Everybody takes the quiz but the way I record the grade is only the good grades. If you don’t get a B or better then I don’t record it.
- I also try to give opportunities and assess them in different types of ways and give them the opportunity to, if they blew something, ... I give them the opportunity to make up points or get bonuses.
- When we do tests and quizzes I’ll divide them into sections so each of my learners will have at least one area on the quiz or test where they’ll be able to shine.
- You want to have a variety of activities, because they may shine in one and not in the other.
- I always found that the more ways you can assess children and the more grades you have for them, if they have failed a couple of tests or couple of projects or missed a couple of homeworks, you can take that out and still keep their grade up because you have a large quantity.
- I never tell them this but when I do that average I’ll add an extra 5 points to everybody, and as far as a borderline is concerned, most of the time I round up.
- When I have a kid that is between a D and an F, I go back and really scrutinize [so see] if there is something maybe I did or left out or didn’t ... I’ll go back and re-grade something.
- I always tell them that they are here to succeed, not to fail, and my classroom’s designed for them to succeed.

Promoting Student Understanding

An important component of the philosophy of teaching/learning is to gauge student progress by using assessment to check for student understanding. This came up many times in the

responses of the teachers. They are very concerned about getting students to the point where they truly comprehend and understand, not merely memorize.

- It's not only me lecturing and then they soak it in and then regurgitate back to me. I always tell them that I could train a monkey to do that if I give them enough bananas. That's not what education is, that's not my goal here. My goal is that you can understand it, so you need to participate.
- And to me that's how you approach math. It's not memorizing because first of all you're not going to retain it and that's not going to help you.
- You want to know, what have they really learned or can they apply it ... to get a more realistic grade of what the student really does know about the material.
- So we have to read it in class [Macbeth] ... then I say, 'alright, this is what I'm going to test you on ... I give them samples ... then I decide who gets what kind of quiz. It's a lot of work, but the thing is it makes them more successful.
- The assessments where students actually have to show me some work or write about are most valuable for informing me about how much students know. Because it's then that you know that they understood every process that tells you more about a student than just grading a sheet of answers.
- Big tests and essays best because this is where higher levels of thinking come in.
- I go back to the ultimate, I don't care how I get them there, I want them to learn it. And if it means I will give you 2 more points for this if you go back and fix it and get it right ... if I have to dangle that carrot to get them three ... I'll get them there.

Accommodating Individual Differences Among Students

Another aspect of assessment that appeared linked to philosophy was varying assessments to accommodate individual styles. Most of the teachers made efforts to provide varied assessment that meet a variety of learning styles. This is, again, part of a larger philosophy of education – that individual differences among students are important and need to be considered – in all aspects of teaching, including assessment and grading. What teachers do is essentially modify assessments on the basis of student characteristics. Following are some illustrative quotes:

- I think it'll go back to the goal I have: try to meet the needs, interests, capabilities of the children. If you don't have a variety of things, you're not really focusing on what that child's ability is.

- I will have at least two versions of any given test or assessment ... there are always three versions and most often four depending on who I have in my class ... and I think it is very important to do it ... I always have several ways [of assessing].
- Some students do better on paper than they do orally, some students do terrible on paper but you know that they know it so you have to come up with a way to say show me what you know. My philosophy is I'm trying to get them to show me what they know, not trick them into showing me what they don't know.
- I feel that kids learn in all different ways, and they have different ways of showing it.
- They [assessments] change based on the needs or capabilities of the students. I'll make some tests easier and some harder, depending on the ability level of the student.
- We tend to always have a fairly big disparity in student ability ... there's no use in teaching things that people already know, and there's no use to teaching way over people's heads.
- The types of quizzes do vary .. I try to accommodate. I wouldn't put as many formal proofs on their quizzes or tests as I would with an honors class.
- It's really, really important that you know the kids individually as people and you have to know their stories.

Student Engagement and Motivation

Teachers clearly indicated that it is imperative for students to be actively engaged in learning, and, hopefully, motivated to do their best work. This engagement and motivation is seen as critical to the learning process. Consequently, teachers base their assessment and grading practices decisions on what will result in the greatest amount of student engagement and motivation. As will be pointed out later, this results in using many constructed response assessment items and a fairly heavy emphasis on homework. Here are some quotes from teachers that illustrate the importance of engagement and motivation:

- If you really want a student to learn, the student has to be actively engaged ... and doing group work ... I find that works best.
- Everybody has to be involved in this, not just those who look like they are falling asleep, but everybody ... we'll continue until everybody has their chance.
- The reason I do it [use a mix of assessments] is because I want my children to be task oriented, I want them to be responsible for every assignment a teacher gives. I don't want them to think they can skip an assignment.
- Students learn more when students are actively engaged. Daily grades are based on participation in groups.
- Students need to be task oriented and organized.

- It's worth a few extra points in their grade because it means that everybody in the whole school triangle [student, parents, teacher] is involved in their education.
- Use essays to get them engaged, to motivate them.
- I use a goody jar ... it's not really assessment ... but really helps me in my assessment, especially with kids with low motivation.
- I make them do something. Make them learn and see that they have to put forth some effort. They need to know that they have earned what they've got.

External Factors

External factors are influences that originate outside of the classroom. They are not under the control of the teacher, but still impact the nature of classroom assessments and grading practices.

Three major external factors were identified: Standards of Learning (SOL) and SOL tests, district grading policies, and parents.

SOL and SOL Tests

This category was included on the interview guide rather than inductively derived from the data. Since it is clear that the SOL and SOL tests have had a great influence on teachers in general, the intent here was to focus the teachers' thinking on how the SOL tests affected their classroom assessment and grading practices. The teacher comments indicated that the SOL were, in fact, impacting their classroom assessments and grading. This was typically not a radical or far-reaching influence. Rather, the SOL tests provided an "external" reason to modify their classroom assessments so that they covered more of the SOL, and, to a lesser extent, so that they used the multiple choice question format to a greater extent. Using more multiple choice questions would better prepare students for the SOL tests. Teachers seemed to feel resigned to making these changes and seemed to suggest that without the SOL tests they probably would not have made the changes. The comments below capture these perceptions of the teachers. In the

comments it is evident that there is tension between what the SOL and SOL tests suggest should be assessed, and how, and the assessment approaches of teachers based on their own beliefs.

- I use teacher made tests because I feel like I'm the one who has taught it and I know what I'm looking for. But as the same time I'm going to make it up according to the SOL that we have to follow.
- What they did was they [SOL tests] defined it [classroom assessments] ... in math most of them are not multiple choice tests, but I give them more multiple choice so they can get used to it ... with a multiple choice test I don't think you get an accurate evaluation of the students' knowledge.
- I think the teacher has to teach both for the SOL test, which is a necessity today, but you can't forget a lot of other things ... you have to have a balance of both.
- I do it sporadically so it's familiar to them, but it's not my general way of doing things.
- As far as changing my grading practices, probably it will change my assessments somewhat. I've got to make myself do more multiple choice questions.
- We're testing more often, we're giving samples of standardized tests, multiple choice tests.
- On my wall over there is SOLs and I know exactly what's got to be on that test because we have practice SOL tests that we've been given to preassess the students. So assessments are beginning to drive my lesson plans.
- Assessments are now not just to know (what students know), but to prepare them to have test taking skills.
- It is a good thing in terms of having them ready for standardized testing.

The following comments show how some teachers have changed their classroom assessments to conform to the SOL, but do not believe this is in the best interest of the students. The impact is undesirable because it means content not on the SOL tests is much less likely to be assessed or emphasized. Here assessments are driving instruction, and to the extent that the classroom assessments are influenced by the SOL and SOL tests, the greater this external influence is on teacher practices. Conflict also arises between assessments that teachers believe give them greater understanding of student knowledge, typically constructed response items, and multiple choice tests, which are viewed as limited in what they tell the teacher.

- I think you're doing the children such a horrible disservice when you teach the

SOL tests because you leave out so much wonderful stuff that some of these children will never get anywhere else.

- I am opposed to the SOL testing, it just doesn't leave any room for individualization on the part of the teacher.
- They [multiple choice tests] don't always tell me what I want to know.
- I like constructed answers because you can go back and show the child where they messed up, I guess multiple choice measures if the child knows or that the child is a good test taker.
- You know frankly, with few exceptions, if we didn't have to get ready for standardized tests I'd probably seldom use multiple choice as a format.
- I'll tell you that this year has been very different from the years past because I always gave tests and quizzes that required show me the work. I didn't give multiple choice tests. Because the students now are required to take these SOL tests, now I'm mixing ... some multiple choice and some show me the work ... yea, outside forces control me.
- This year with the SOL coming in a lot of revision is needed.

District Grading Policies

Each teacher was asked about the effect of division grading policies on their grading practices. While each division has such a policy, it was evident that teachers use such policies only in a very general way, and that their own approaches and preferences are much more important. This contributes to greater diversity of grading practices among teachers. In some cases teachers completely ignored division policies.

- Fifty percent or less [is driven by the school or division policy].
- It's my decision as to how I interpret school policy.
- I am somewhat compelled to go with our numerical system with the county.
- I go with my own judgment also; a little bit of both to be totally honest.
- We got one [division policy] this year. I was furious about it. I'm finally getting some things right after 30 years, and they told me I couldn't do things ... it's not the grading, it's the process of learning.
- I think they [division] are more concerned with you having enough grades every 9 weeks.

Parents

This category is one that was derived inductively from the data. There was little parental influence on the nature of assessments, but clearly teachers were influenced by parents in the grading system they use. Teachers want to be able to meet with parents and provide reasonable

explanations for the grades they have given. The most important factor is having sufficient justification for grades to avoid parental conflicts. For some teachers this meant using very specific, "objective" scores and averaging, while for most teachers there needed to be a sufficient number of grades to show clear patterns.

- To me the calculator is the deciding factor ... I can sit down with any parent or any kid ... it makes my life easy, because I punch the buttons, hit the equal sign, and there it is.
- When you're going to show Johnny's mom why he got a B, you won't have a lot of reasons to show his mom if you only have one test out a nine weeks.
- Parents were a bit more prone to suing teachers ... as a result we had to develop our own objective ways of assessing students' academic performance.
- I don't give A students a C without their parents knowing ... challenges come from a lack of communication between the parents and teachers ... I bring parents into the process on the first day ... talk to my parents all the time.
- I'm very up front with them on how I do report card grades ... I tell them that when they get report cards they will be based on objectives and how well students met them.
- Pretty iron clad, does not leave any room for any subjective evaluation at all. It saves a whole lot of arguments with parents.

Teacher Decision-Making

Teachers were asked repeatedly to provide a rationale or justification for *why* they made assessment and grading practices decisions (e.g., why the types of assessments used, why specific factors used in grading). Overall there was great difficulty and some uneasiness in the responses of the teachers. It was something of a struggle for them to explain, particularly if they had been teaching for many years. One general assumption seemed to provide a foundation or basis for their rationales. It was apparent from the interviews that teacher decision-making is a highly individualized, idiosyncratic process. Thus, no two teachers were alike, and the comments suggested that they believed they should not use the same assessment and grading practices as other teachers. Furthermore, with some probing, five additional factors emerged as significant in identifying the reasons teachers give for their assessment and grading decisions:

the nature of learning objectives, using a wide range of practices, the importance of constructed response assessments, professional experience, and homework.

Nature of Learning Objectives

Many of the teachers indicated that the nature of the learning objective would determine the choice of assessment method. Simple recall knowledge, emphasized through drill and memorization, would be assessed using selected response items, such as multiple choice or true/false, while objectives emphasizing thinking skills, such as application and reasoning, would be assessed with constructed response assessments, such as essays and performance assessments.

The following excerpts show the influence of objectives and topics.

- Well it depends on the topic sometimes, for example, I just finished the objective on surface area and volume and in that case I do a lot of handouts and worksheets. With learning definitions I'll do matching, multiple choice type items. So a lot of times the topic will determine the assessment that goes with it.
- The manner that I assess tends to be more related to the subject matter that I am teaching ... with teaching grammar, just multiple choice items type things for that because there is some memorization involved.
- I do some grammar quizzes, we are going to have a grammar test, they're really sporadic because it really has to be revealed in their writing.
- It depended on what were covering as to how we needed to assess them.
- Pop quizzes that they don't know about that I usually try to give them to get an idea of whether or not they understand the material.
- I try to assess at the point that I feel pretty confident that the children understand the material. That's the point at which I assess.

Using A Wide Range of Practices

A second finding was that all the teachers believed that they should use a variety of assessment methods and should use multiple criteria in grading. This may reflect the conflicting influences of internal and external factors, but probably is based mostly on the belief that since no two students learn the same, multiple assessment methods are needed to fairly assess them so

that all students are able to demonstrate what they have learned. It's also consistent with the "pulling for students" belief, use whatever assessment best matches with student styles and strengths to give the best performance. Notice in the following comments how assessment practices are both varied and influenced by the nature of the students.

- It's what I feel like the kids in that particular group, how I'm going to find out what they know in the best way ... some students do better on paper than they do orally, some students do terribly on paper but you know that they know it so you have to come up with a way to say show me what you know. My philosophy is I'm trying to get them to show me what they know.
- I use a little of everything.
- Day to day, observation, the almighty observation. You're listening, oral presentations, looking at a project, at a test score, it's everything that a child can give you.
- They [assessment practices] change based on the needs or capabilities of the students. I'll make some tests easier and some harder depending on the ability level of the students.
- You have to adjust [assessment practices] to where you are.
- I tend to rotate they types of assessments so that they have a lot of different types.
- I try to give opportunities and assess them in different ways and give them the opportunity to ... if they blew something, in a projects format it seems like its easier to assess ... I give them the opportunity to make up points.

Importance of Constructed Response Assessments

It was clear that the vast majority of teachers preferred constructed response assessments, where students "show their work" (e.g., short answer, essay, performance assessments, demonstrations, exhibitions, portfolios). The teachers indicated that these kinds of assessments give them the best indication of whether students truly understand and can apply what they have learned. This is consistent with the internal belief that assessments should serve instruction by showing what students understand. Caution was indicated in the extent to which objective assessments can provide sufficient evidence that students actually understand as compared to memorization. Here are some illustrative comments:

- Whereby I use rubrics to score a lot of their projects, I also try to do as much hands on as possible.

- Observations, rubrics, anything that will show you a measurement of a child's performance level.
- I like the rubrics. Just because it allows for more creativity on the child's part. It seems like they are giving me more information.
- One project I like is called shop till you drop. They're required to go out and comparison shop, try to find the same items on sale at three different stores ... they have to do different things to get it ... so I use that as a test grade.
- I use written open format evaluation. I occasionally use matching and multiple choice. I do a lot of personal anecdotal evaluations. In other words, live performance tests.
- I might start with a quiz and then if it was still unclear then I would go to a personal one-on-one oral assessment or a task assessment ... a project type of assessment.
- When it comes to science it would tend to be more free form drawing open-ended questions some task oriented things where they have to do something ... the same goes with social studies ... more open-ended assessments too with some hard knowledge types of stuff.
- I have some multiple choice, personally I do not like those, I would rather have free response because then they have to put down exactly what they know.
- I always use rubrics ... and when we do book reports, or any product based outcome, when we're building or making something ... we use the rubric.
- My tests are for the most part essay ... generally really thorough and quite long ... all the kids have to write three essays.
- If I teach geography, the way I would assess that is to give them a blank map.

Professional Experience

Teachers' experiences have evidently had much to do with determining their assessment and grading practices. Whether by trial and error, or by talking with others, it seems that the teachers learned through their own experience which assessments and grading approaches would work best for them and their students. It is as if the practices simply evolved over time. One thing that was absent in their comments was any indication of influence from either initial teacher training or subsequent professional development opportunities. The following comments illustrate the importance of experience:

- I've taught for twenty some years and I guess some of this just evolves over the years.
- I had to figure out what to do. Sometimes you talk with other teachers and find that they are doing different things, but I don't know that I have talked to any other teachers who are doing what I do. It sort of came upon me ... trial and error would be the best answer,

which would put it all in a nutshell ... like a lot of things, once you do it for a long time you sorta get a feel for it.

- Test experience, what was done with me in high school.

Homework

Finally, one common thread among most teachers was the importance of homework. It was indicated that homework, much like quizzes, provide the teacher with an immediate indication of student understanding. Homework also is important to student learning. Most teachers believed that homework was essential practice in the skills. In this sense, then, homework provides an added learning activity and an indication of understanding. Here instruction and assessment are integrated. Here are a few comments indicating the importance of homework:

- I weight [grades] more on homework, say around 40%.
- I found out that if a student starts homework in class ... [put] homework assignments on the board and go over them ... you ought to be able to help them.
- I give what I call a mini quiz every day. That's o the last night's homework.
- I'll teach a lesson, they'll have a homework assignment, they bring it in the next day and we'll check over it and I'll check and see who has it because use that way then you get something wrong on the homework, I expect them to ask questions.

Assessment Practices

Based on beliefs and values, and on external influences, teachers select and implement specific assessment practices. The variety of different types of assessments used reflects teacher beliefs that informal, observational assessments and constructed response assessment are best for gauging student understand on the one hand, while external pressures tend to result in more objective items. As a result, most teachers use the same variety of assessments, individualized to their students and based on their own experience and the nature of the learning objectives. This would include homework, quizzes, tests, performance assessments, and participation. Using different kinds of assessments also allows more students to show their best work. Several

themes concerning assessment practices emerged from the data, including formative assessments, pre-assessments, revisions of assessments, and construction of assessments.

Formative Assessments

From the responses of the teachers it became evident that some assessments are more informative than others. The daily checks and observations, what might be called formative or informal assessment, were clearly most informative for the teachers. This kind of assessment is ongoing and continually informs instructional decision-making. Here are some examples of what they said about the informativeness of assessments:

- My informal assessment is [most informative].
- Daily quizzes. Yes, especially with daily quizzes as a check of previous day and use quiz to go over concepts as needed.
- Ones where students apply what they've learned [are most informative].
- Daily quizzes [are most informative] ... gives daily pulse of learning for the teacher.
- Quiz, graded by both teacher and student [is most informative]. Gives a quick overview of class progress.
- Well, definitely the free response over the multiple choice ... sometimes the group assignments I give ... also listening to them talking in groups.
- Using a rubric with very specific guidelines [is most informative] for me and them [students].
- Probably I can find out a whole lot more in oral. Asking them to explain something to me ... and just watching them on a day-to-day basis.
- Probably the tests. Because like I said, you can pretty much realize that they have been able to master larger chunks of information that's opposed to isolated things.
- Almost always for my purpose it's the writing assignments that I spend in the rubric [that is most informative].
- Probably quizzes on 2 or 3 sections [are most informative].
- Class work assignments [are most informative] because they are one in class under supervision.
- Projects most important, also papers.
- Oral questions used extensively and daily assessments.
- Homework and class participation.
- I develop a rubric for every writing assignment and I show the children that rubric in advance as well. ... it lists the skills I'm looking for ... it's a diagnostic kind of thing ... it gives me the ability to cover things they've learned in the past and weak spots.

Pre-Assessments

Another area that was brought up in the interviews was the nature and use of pre-assessments. These are assessments done prior to instruction or beginning a unit. It was clear that most teachers use some kind of pre-assessment. This was usually in the form of an informal review of current student knowledge, understanding, and skill, done through classroom observation, short quizzes, and through question/answer sessions. However, some teachers actually gave formal pretests. Also, some teachers interpreted “pre-assessment” to mean “expectations,” which they tried not to make. Finally, pre-assessments are affected by subject matter and experience. In highly sequenced subjects, such as secondary mathematics, there is less pre-assessment. The more experience a teacher has, the less likely he or she will use pre-assessments.

Here is a representative sample of comments of teachers regarding pre-assessments:

- I'd rather judge on what they are doing, not on a standardized test.
- I do pre and post-tests three or four times a year ... if you don't do your pretest in the first or second week of school then you can just hang it up.
- If you preassess then you're going to be able to plan better and you find out student needs.
- I use writing as pretty much a gauge to find out what the students are lacking in and what they need to know.
- I also give them a pre-assessment.
- I always assess early on to see what people know so that I could split groups basically and remediate or accelerate as needed. I am a real stickler to assessing only to find out what people know.
- What I will do is I will kind of explore their knowledge so to speak. I might start with a quiz and then if it was still unclear then I would do a personal one-on-one oral assessment.
- At the beginning of the year I'll give them a math test o the SOLs and let them see what they couldn't do.
- Sometimes I will use a diagnostic test ... some of the kids you think are very articulate but that doesn't mean they know some of it.
- Yes, with daily quizzes, especially with advanced students who need to move ahead of the rest of the class.

- Informally, usually in beginning a unit.
- Not so much in algebra ... mostly new material ... find pre-assessing discouraging because so much is new.
- Pretest to find location that class is in to start the year.
- Not as much as I probably should ... I used to give a pre-assessment of grammar just to see where they were ... they were all over the place.
- Sometimes we do a little bit of that [pre-assessment]. I usually have a pretest, maybe the first day to kind of see where they stand. I used to do it more than I do now.
- We have a little survey we do.
- Sometimes. There are only four math teachers here and I know the 6th grade teacher and what she covers ... I only teach those things they don't know.
- Yea, because at the beginning of the year for the math class there is a lot of review.
- No, because they have been pre-assessed because of the placement into the pre-algebra group.
- Yes, informal, with short quizzes or examples.
- Yes, informally with questions and answers at the board.

Revisions of Assessments

Teachers were asked to comment on the extent to which they revise assessments, when the revisions would be done, and the nature of the revisions. Almost all of the teachers indicated that tests and other assessments are revised both from year to year and from one week to the next as the testing date approaches. This constant revision process is done because students change from one year to the next, because the content of what is covered changes somewhat, and because in each class there are special circumstances that affect what should be tested. This further supports the teachers' need to adjust assessments to individual differences of students and to the objectives being covered, all in the goal of pursuing increased student learning. It also points to a significant time commitment for teachers.

- Every year it's different ... I don't think I've ever used the same test two years in a row on anything.
- The only thing that has stayed the same is the spelling quizzes ... I change everything else ... I keep a copy of it so I can see the change.
- You modify everything. If everybody fails the test then I modify it because I've done something wrong ... I try to write and revise tests students take within the next two days.

- Why are they revised? Because the results that were found on previous tests were not satisfactory, did not show student performance.
- I usually change them pretty much each year.
- I rewrite them every year, maybe not entirely, I'll use some parts.
- When I grade a set of papers and there is something there that the children are not understanding, I go back and revise the assessment ... so there's a constant revision process going on.
- It [revision] gets done on a regular basis ... this year I've modified almost all of my tests.
- If I find an old quiz, it just doesn't work out for them [students] ... it's the wrong thing. They're not there yet or they're way behind it or what have you ... it [using old tests again] just doesn't work.
- The weekly tests, you're constantly changing based on the needs of the class.
- Yes, I do go back and modify them [tests], I don't just pull it out of the cabinet and give it to them.
- I don't even look from one year to the next to see, I always rewrite them. I know from last year to this year they are a whole lot different. I don't just recycle.

Construction of Assessments

Teachers may use assessments they themselves construct, they may use tests provided by publishers or school divisions, or they may use some combination of these, each influencing the other. Overall, these teachers clearly rely most on teacher-made tests, ones they construct. Publisher's tests are not widely used because they do not address local contextual factors such as what was covered and the characteristics of the students. Note how the following quotes emphasize the importance of teacher-made assessments:

- Over the whole year generally, I do teacher-made tests.
- I use teacher-made tests because I feel like I'm the one who has taught it and I know what I'm looking for.
- Sometimes I will pull questions from a pre-made test but I don't generally like to give an entirely text book made test ... I don't tend to teach things like they are presented in the book, so I make them [tests] up.
- Some of them I create myself, some of them are from the text. When I take them from the text, I very rarely give the whole thing, I usually do bits and pieces and kind of paste and put together.

Grading Practices

Grading practices represent an interesting mix of results from assessments and deciding how to weight different factors different amounts. In addition, there clearly are external factors, such as division grading policies and parents, as well as teacher beliefs about motivation and engagement that influence the practices. What results is, like the nature of assessments, individualized approaches that take these considerations into account with the types of students in the class. In discussing grades several factors emerged as significant, including grading policy, borderline grades, how effort is handled, how extra credit is handled, grade distributions, and how zeros are handled.

Grading Policy

Regardless of division or school policy, teachers have their own grading policy. And it seems that most teachers have unique or idiosyncratic procedures. However there are some common elements. For one, all teachers obtain many grades from primarily four sources, homework, quizzes, tests, and projects or papers. Some teachers also utilize participation, class work, or some other indicator of effort. Interestingly, tests typically do not account for more than 30% of the final grade. Teachers indicated that they use a criterion-referenced approach to grading rather than a norm-referenced approach, and typically would use a total point system that provides percentages consistent with division guidelines. An interesting issue is whether the teacher uses students in the class or grade level objectives as a frame of reference for giving grades. That is, it is possible for students to receive As if they learn a lot, or receive Cs for the same level of performance if what they have achieved is below grade level.

- I rely on tests only 30%; class work 65%.

- I'm not a believer in having a bell shaped curve for grades ... if in a class nobody's trying and I only have one or two Bs and the rest are Fs, that's exactly what the assessments are going to be.
- If I have students who are working on the first grade level they necessarily get Ds and Fs on the report card where I'm grading 3rd grade objectives.
- They give me objectives ... I always to way , way beyond those.
- They have 13 grades, I drop the 3 lowest ... I figure everybody has an off day.
- I'll have more grades than I know what to do with.
- I give quizzes and tests and I work on a point-total system
- I would break the quizzes, tests, and homework into a percentage grade.

Borderline Grades

Every teacher faces decisions concerning grades that are borderline, just between two letter grades, or very close to a higher grade. Teachers in this sample indicated that in these situations they want to be able to give students the benefit of the doubt (pulling for students), and typically use non-achievement factors for making their decision, such as effort and participation, or use extra work or extra credit. This reflects the teachers' desire for students to be as successful as possible and to obtain the highest grade possible. It is usually a subjective judgment by the teacher. Here are some illustrations of what teachers do with borderline grades:

- I will suggest that maybe they do something extra, which could be a project ... I tutor with them ... I'll give them make-up work because usually they don't even ask for make-up work they missed. It depends on the situation, but I do what I can to try to help them over the hump.
- Borderline comes down to effort.
- Borderline, effort is the key ... can make up zeros or use extra credit.
- If they come in and say they got a 60 the first time and they come in and get a 85, then I'll up that to a 75 or something.
- An F is a 63, those kids get 60. I will pass them especially if they've really showed me the effort ... if I know they're really trying and I mean, genuinely, then I will pass them.
- If I'm within a point or a point and a half of the next letter grade, I look at the child and do I feel the child has made an effort?
- Then I generally think of their effort, whether I feel they've really tried and whether they've turned in all their work. If they didn't turn it all in and it's borderline, I don't give it to them ... if they tried to make an effort to improve, I won't give them an F; if

they didn't do their work and they've been absent quite a bit, then they're gonna get what they deserve.

- When it's borderline, how hard has the child worked in the year? And I will be honest with them, it it's a 63.5, I'm going to bring it up to 64.
- Borderline, most of the time I round up ... I'll give extra points to someone who really works hard.
- Reserve A for performance, B for effort is possible.

There were a few teachers who clearly did not want to use subjective criteria for borderline situations:

- Frankly I tell them that when they get report cards they will be abased oh objectives and how well people meet them. How can I grade on effort?
- The calculator decides [borderline cases] ... to me, I round up half a point ... I try to set up the system where I don't have to make evaluative judgments.

Student Effort

One of the most varied practices in grading students is concerned with how the teacher recognizes and handles student effort. From one standpoint, most teachers use effort to some extent in deciding borderline cases, giving a student who tries hard the higher grade. Many teachers view effort as enabling achievement or as part of achievement, so that it becomes an important contributor to determining grades. Some teachers do not use effort at all, relying instead solely on the quality of student performance. Many teachers think of homework as a proxy for effort. The following quotes show how different teachers have different ideas about how to handle effort.

- At this level you have to take into consideration effort but it can't be to the exclusion of performance because it's a fine line.
- I have one child I think is getting a D and she had worked like a dog and so we really just bumped her up to a C.
- As far as an effort grade, I don't really believe in effort grades but, well, homework is a good example, I give an effort grade for having homework everyday.
- Most projects there is usually a window where I'm grading effort. I can tell that some have been working really hard and I'm going to give them the benefit of the doubt

...there is one girl that tries really hard and all she can get is a high F, and I give her a D every time. I will not fail that girl.

- I want to see that there were sincere efforts. When I look and see that a child's missed eight out of ten homework assignments ... he decided just to sit there and not do them ... that's what I measure as not sincere efforts.
- I put effort in their class participation grade. Some students sit there and don't say a word. I factor in not only their actual class participation, but also their effort, what I perceive is effort.
- So to me, conduct and behavior and attendance is very, very important in assessing that final grade.
- It [effort] only comes into play in that test and essay realm and then in the end result ... if Johnny probably deserved to have an 83, I would maybe for that effort give him above for the grade ... and on the other side of that coin, I would maybe not bump her down.

Extra Credit

Teachers were asked how they incorporated extra credit in grading students, if at all. Most teachers do use extra credit, primarily as a way of boosting the grades of students who may be borderline or receiving a low grade (pulling for students). There are many different ways extra credit is used. Some teachers make it relatively easy for students and have an informal set of guidelines, while other teachers believe students must clearly earn the extra credit with additional effort. Another variable with extra credit is whether students know about it and can plan for it, or whether the teacher simply awards extra credit as a surprise. Both approaches are used. Many teachers offer ways to earn extra points as extra credit. Some comments of teachers about extra credit are the following:

- I tell them they can have extra credit when they have done what they have supposed to do for credit. Make them learn and see that they have to put forth some effort. I think too many kids get by today with not earning what they get, and that's an important lesson.
- If somebody does extra credit and it doesn't indicate better performance, then no I'm not going to give them anything. There's no like free points. I always retest everybody who gets Ds and Fs, and I'll throw out the old one. I will always give people a chance to improve, why not?
- The things that motivate my kids is they'll put so much more effort into the extra credit than they will the regular work. They love to see 75 + 10.

- If it was a particularly hard homework question, the ones who got that when I go around ... I give them extra credit. They never knew that part ahead of time.
- I rather them do the assignment themselves rather than give extra credit. But what I do is offer bonus points, which, I guess is almost the same. For example, just things like taking home papers to get signed if they bring them back.
- It's worth a few extra points if they're willing to show things to their parents to keep them abreast of what's going on in the classroom.
- I don't give extra credit. I tell students that you earn your grade ... you don't come in at the last minute and ask for a bail out ... [but] we do have extra assignments that are optional that you can do to earn extra points.
- They get two make-up assignments; that's the extra credit.
- Sometimes I'll give them an extra credit problem or a project or something like that.
- I have one class when they have to bring their report cards back signed they'll get 3 or 4 points.

Handling Zeros

A vexing problem in grading students is how to handle zeros. Our teachers reported a variety of ways that zeros are used. Teachers generally understand the devastating effect a zero can have on grades, and most teachers try to accommodate students by providing opportunities to remove zeros (pulling for students). Some teachers use zeros for motivation. Generally a zero is intended for no work at all, not for receiving an F. Like other assessment and grading practices, zeros are handled in ways that make most sense for individual students, despite the presence of a single policy.

- A zero means you didn't do the project at all ... an F means you did the work and you deserve some credit. For the most part I try not to let the kids get a D or F, I have what you call do-overs.
- If they [students] just got one zero, I mean I'm lenient enough. They are not going to figure out the percentages anyway, so I can fix it then.
- I put the zero in at the end of the nine weeks if they just haven't turned anything in ... I try to make sure they have an opportunity to make it up. I know a zero will kill their grade and they don't understand that.
- I have a lot of grades, so one zero does not make a great deal of difference ... it's all done in percentages so at the beginning it has a heavy effect ... I do not give them a chance to make it up.

- It [zero] counts as a regular grade. One of the things I discuss when we are covering means is that every zero counts, don't miss assignments and think you're getting over, you're not.
- Oh, I record them [zeros] to start with, but I don't know. Maybe this is the dedicated teacher's syndrome or whatever. I'll chase the kid around for a long time so that I can get a few points. I have a child now who is absolutely and A+ student. She hasn't turned in her last writing assignment ... so it's dropped her A+ to a C- ... I've hounded her every single day.
- I don't give a zero ... it's murder for a child to make up. There are people that give 0's and it just turns the kids right off.
- I cannot change a grade ... the zeros stay there ... the zero stays if they don't make it up ... there's a lot of stuff I want to broadcast, but I just can't turn them down when a kid comes to me.
- If it is a graded assignment then yes, I consider it a zero, but I offer them an opportunity to go back and do them. It's the learning that I'm most interested in, not the penalizing.
- It's very straightforward. They are just average in and if there are mitigating circumstances then I would take that into consideration.

Conclusions

The results of these analyses indicates that teachers have a lot to say about their idiosyncratic assessment and grading practices. It appears that teachers are constantly striving to reach a reasonable balance between their beliefs about education and learning on the one hand, with the pressures exerted by external factors. This constant state of tension may help explain why teachers view assessment and grading as a fluid set of principles that change to some extent each year. Together, these influences converge on the actual process of making assessment and grading decision, which result in turn in the development and implementation of assessment and grading practices. Because of the interplay between the teachers' beliefs, external factors, and student characteristics, a great amount of variety in classroom assessment and grading is evident.

Important teacher beliefs that influence decision-making include a larger philosophy of teaching and learning, wanting students to succeed, accommodating individual differences among students, engaging and motivating students to learn, and promoting student understanding

and mastery. These beliefs converge on getting students, in whatever ways are necessary, to be involved in learning, giving effort, and ultimately demonstrating successful performance.

Important external forces include the SOL and SOL tests, district grading policies, and parents. Clearly, the most important external factor are the SOL and SOL tests. These externally mandated high stakes tests have put pressure on teachers to modify their assessment practices to accommodate the SOL and the format of the SOL tests.

One impression is the strong sense of ownership teachers have for their assessment and grading practices. It is almost as if there is a sense of pride and ownership that the practices are unique and that they have a good rationale for them. It also seems that assessment and grading is largely a private business, not readily talked about very much with other teachers. Clearly, assessment and grading practices fit within a larger philosophy of student learning, and clearly teachers are very interested in and committed to enhancing the learning of each student. They want students to learn. So it follows that they want assessment and grading practices to enhance student learning, not simply document student performance.

Assessment practices that emerged from the interviews stressed the wide variety of assessments used for different purposes, and the need for variety to accommodate student learning styles. Formative assessments are used constantly during instruction to inform teaching decisions. Pre-assessments are sometimes used prior to instruction to gauge current student knowledge. Revisions are made continuously by teachers, and teachers, in the main, construct the assessments they use with their students.

Grading practices are very idiosyncratic. Teachers adopt their own grading policy, with little regard for standardization with other teachers. Most teachers use effort as a determining factor in borderline grades, and in general believe that student effort is a good proxy for student

achievement. Extra credit is used to help students obtain a higher grade. There is great variety in how zeros are handled.

An important finding from these data is that classroom assessment and grading are integrated with instruction. Most teachers see assessment and grading as extensions of instruction that have important consequences for student engagement and motivation. Thus, teachers' decision-making is heavily influenced by thinking about how the assessments will enhance student learning. Teachers believe that learning is best assessed with multiple assessments, using different formats. They also believe that informal or formative, and constructed-response assessments provide the best information to judge student understanding.

Our goal in this study was to "get inside the head" of teachers to find out what influences their decision-making concerning assessment and grading practices. We have learned that decisions are made on the basis of how the assessment or grading procedure will affect student learning and motivation, and, at the same time, respond to external pressures. In this balancing act each teacher has his or her own solution, one that is constantly changing with each new group of students.

Implications

The results of this study suggest several implications. First, given that teachers clearly "pull" for student success and use many different practices that help student succeed, it may be helpful to ask if teachers are "coddling" students, making it so easy to obtain passing and even high grades that students are getting a false sense of their own level of understanding and performance. In other words, is the desire of the teacher to see student "success" so strong that it

promotes assessment and grading practices that students can obtain good grades without really knowing the content or being able to demonstrate the skill?

Second, what are the results of emphasizing effort as much as teachers do in grading students?

Research on student motivation and attributions for success (reasons students give for their success) suggest that while an emphasis on effort is positive for motivation because effort is a controllable, internal factor, it may be counterproductive for some low performing students because they may develop a belief that they can be rewarded for effort and not mastery of the content or skills. This may also give students a false sense of their competence. Furthermore, too great an emphasis on effort may reduce attributions to ability, which are more stable. On the other hand, this emphasis on effort at least teaches students the importance of engagement and involvement and the need for this involvement to be successful.

A third implication is concerned with the skills teachers have to construct and revise classroom assessments. It is clear from the literature, and the results from Phase 1 of this research effort, that teachers may not have the knowledge and skills that are needed to effectively construct and revise assessments. With the popularity of new types of assessments, such as performance and portfolio assessments, teacher skills in assessment may be thinned even further. It may be helpful to systematically evaluate teachers' assessment skills and provide professional development where needed.

A fourth implication of these findings is the potential effect of external pressures on teacher professionalism. The influence of the SOL and SOL tests is undeniable, and seems to be directed at something that is very important to teachers' sense of what it means to be an effective teacher. Teachers desire autonomy and need to adapt instruction and assessment to their personal styles and to the needs of individual students. Teachers do not appreciate standardization of practices

that minimize these dimensions of being a teacher, and the SOL and SOL tests have had such a standardization effect. The question is whether this, in fact, is affecting teachers' sense of professionalism, and if so, what impact this has on teacher morale and motivation. In addition, it may be that in Virginia, at this time, external pressures are particularly influential given the current situation with the SOL testing.

A fifth implication concerns teacher training and teacher induction. What do these data suggest with respect to how teachers are trained? How important is it for teachers to have a fully developed philosophy of teaching and learning so that assessment and grading practices can be based on this philosophy? What is being done in teacher training to help teachers become competent in the variety of assessment methods that are typically used, as well as how to integrate external pressures with personal beliefs and district grading policies? In the induction of beginning teachers it may be valuable to examine their assessment and grading practices to see if they are consistent with philosophy of teaching and learning and other beliefs. For example, if a strong value in teaching in maximizing the learning of each student, what adjustments in assessments are made to accommodate individual differences among students?

It is clear that teachers spend a great deal of time with assessment and grading, and that they see these tasks as integral to the teaching/learning process. This research helps to show how teachers make assessment and grading decisions, pointing to tension teachers feel when internal beliefs and values conflict with external pressures and demands. This understanding will hopefully suggest positive actions that can improve assessment and grading practices.

References

- Airasian, P. W. (1997). Classroom assessment (3rd Edition). NY: McGraw-Hill.
- Airasian, P. W. (1984). Classroom assessment and educational improvement. Paper presented at the conference Classroom Assessment: A Key to Educational Excellence, Northwest Regional Educational Laboratory, Portland, Oregon.
- American Federation of Teachers, National Council on Measurement in Education, and National Education Association (AFT, NCME, NEA). (1990) Standards for teacher competence in educational measurement. Washington, DC: Author.
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. Journal of Educational Psychology, 84, 261-271.
- Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. Applied Measurement in Education, 10, 161-180.
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. Applied Measurement in Education, 7, 279-301.
- Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. Journal of Educational Measurement, 30, 123-142.
- Brookhart, S. M. (1991). Grading practices and validity. Educational Measurement: Issues and Practice, 10, 35-36.
- Carter, K. (1984). Do teachers understand the principles for writing tests? Journal of Teacher Education, 35, 57-60.
- Cizek, G. J. (1997). Learning, achievement, and assessment: Constructs at a crossroads. In G. D. Phye, Editor, Handbook of classroom assessment, NY: Academic Press.
- Cizek, G. J., Fitzgerald, Shawn M., & Rachor, Robert E. (1995). Teachers' Assessment Practices: Preparation, Isolation and the Kitchen Sink. Educational Assessment, 3(2), 159-179.
- Cross, L. H., & Frary, R. B. (1996). Hodgepodge grading: Endorsed by students and teachers alike. Paper presented at the annual meeting of the National Council on Measurement in Education, New York.
- Fleming, M. & Chambers, B. (1983). Teacher-made tests: Windows on the classroom. In W. E. Hathaway, ed., Testing in the schools. New directions for testing and measurement. San Francisco: Jossey-Bass.

- Frary, R.B., Cross, L.H. & Weber, L.J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects: Implications for instruction in measurement. Educational Measurement: Issues and Practice 12(3), 23-30.
- Friedman, S. J. & Manley, M. (1991). Grading practices in the secondary school: Perceptions of the stakeholders. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago.
- Haertel, E., Ferrara, S., Korpi, M., & Prescott, B. (1984). Testing in secondary schools: Student perspectives. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- McMillan, J. H. (1997). Classroom assessment: Principles and practice for effective instruction. Boston: Allyn & Bacon.
- Messick, S. (1989). Validity. In R. L. Linn (ed.), Educational measurement (3rd Edition). Washington, DC: American Council on Education and National Council on Measurement in Education.
- Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. Review of Educational Research, 62, 229-258.
- Phye, G. D. (1997). Classroom assessment: A multidimensional perspective. In G. D. Phye, Editor, Handbook of classroom assessment, NY: Academic Press.
- Plake, B. S., & Impara, J. C. (1997). Teacher assessment literacy: What do teachers know about assessment? In G. D. Phye, ed., Handbook of classroom assessment, NY: Academic Press.
- Plake, B. S., & Impara, J. C. (1993). Assessment competencies of teachers: A national survey. Educational Measurement: Issues and Practice, 12, 10-25.
- Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D. H. Shunk & J. L. Meece (Eds.), Student perceptions in the classroom (pp. 149-183). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Pintrich, P. R., & Schunk, D. H. (1996). Motivation in education: Theory, research, and applications, Englewood Cliffs, NJ: Prentice-Hall.
- Popham, W. J. (1995). Classroom assessment: What teachers need to know, Boston: Allyn & Bacon.
- Schafer, W. D. (1993). Assessment in teacher education. Theory into Practice, 32, 118-126.

Schunk, D. H. (1994). Self-regulation of self-efficacy and attributions in academic settings. In D. H. Schunk & B. J. Zimmerman (Eds.), Self-regulation of learning and performance: Issues and educational applications (pp. 75-99). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Shulman, L. S. (1980). Test design: A view from practice. In E. L. Baker and E. S. Quellmalz, (Eds.), Educational testing and evaluation. Los Angeles: Sage.

Stiggins, R. J. (1997). Student-centered classroom assessment. Upper Saddle River, NJ: Prentice-Hall.

Stiggins, R. J. & Conklin, N. F. (1992). In teacher's hands: Investigating the practices of classroom assessment, Albany, NY: State University of New York.

Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school: Building a research agenda. Educational Measurement: Issues and Practice, 8, 5-14.

Stipek, D. (1998). Motivation to learn: From theory to practice. 3rd ed., Boston: Allyn & Bacon.

Tittle, C. K. (1994). Toward an educational psychology of assessment for teaching and learning: Theories, contexts, and validation arguments. Educational Psychologist, 29, 149-162.

Truog, A. L. & Friedman, S. J. (1996). Evaluating high school teachers' written grading policies from a measurement perspective. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New York.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. Psychological Review, 92(4), 548-73.

Appendix A
Interview Protocol

Topic Guide

2/1/98

Classroom Assessment and Grading Teacher Interview

Directions: The purpose of this Interview Topic Guide is to provide a protocol for asking questions to elicit teacher responses concerning their classroom assessment and grading practices. The general purpose of the interview is to obtain a thorough understanding of **why** teachers use certain assessment and grading practices; **their reasons and decision-making** concerning these practices. It is important for the teachers to be informed that their responses are completely confidential; and they should be encouraged to be as honest as possible. Use the following questions as a guide and make notes of responses in the space provided. After asking the first two questions, feel free to use whatever order seems best in asking the questions, and use prompts as needed.

Begin the interview with “small talk” and other conversation to put the teacher at ease and create a comfortable environment. Define ‘classroom assessment’ as policies, techniques, and procedures used to measure, interpret, and use information to make decisions about what students know and can do on units, chapters, and other major learning goals (e.g., tests, quizzes, homework, papers, reports, etc.).

1. To begin our interview, I'd like you to select a single class you are currently teaching that would be an example of the **most typical class** you usually teach. Then I would like you to answer all of the following questions based on this class. Please describe the class for me, with respect to:

Grade level: _____
Subject: _____
Ability level: _____
Class size: _____

Also, how many years have you taught at this level? _____

2. Briefly, please describe the kinds of assessments that you use for this class.

3. **When** do you decide which assessments to use?

Probe: Before the class begins or during the class?

4. **Why** do you use these particular kinds of assessments?

Probes: Do you ever change assessments? If so, why?
 Do you use them because that's what other teachers use?
 Do you use them because it is what is suggested in instructional materials?
 Do you use them because they motivate students?
 Do you use them because of tradition?
 Do you use them to provide an "objective" record of student performance?
 Do you depend on the school's policy, or is it your decision?

5. Which of your assessments do you find to be most valuable for informing you about how much students know and can do? Why?

Probe: Do you use tests already prepared by the department or publisher?

6. How do you think classroom assessments, like papers, tests, and other assignments affect student motivation? What kinds of assessments seem to motivate students more than other kinds? Please explain why.

Probe: Are your expectations communicated the difficulty of the tests?
 How is student effort assessed?

7. Briefly, what is your grading policy, and **how did you come to decide** what it would be?

Probes: How do you incorporate student effort?
 How do you handle borderline grades?
 How do you use extra credit?
 How do you handle zeros?

8. What is your typical distribution (or spread) of grades given?

Probes: How did you come to believe that that distribution was appropriate?
 Do you ever talk to other teachers about grade distributions?

9. Do you pre-assess students, either formally or informally, to determine their strengths and weaknesses?

Probes: If so, how and how often?

If not, how do you gauge student knowledge and skills prior to instruction?
Is pre-assessment used to better plan instruction to meet student needs?

10. Do your lesson plans determine the assessments you use, or do your assessments dictate your lesson plans?

Probe: Do you use test results to re-instruct students on weaknesses demonstrated through their performances?

11. Do you modify assessments on the basis of student characteristics?

Probes: Do you offer different tests to students at different ability levels?
Different types of assessments?
Make modifications?

12. When do you typically write and revise tests that students take?

Probes: How often are major tests revised?
To what extent are they revised?
Why are the tests revised?

13. What kind of assessments, either formal or informal, do you use day to day to inform you about how much students know and how much progress they are making?

14. How do you give feedback to students when returning an assignment or test?

Probes: Is it done individually or as a group?
Why do you use this kind of feedback and not use other kinds?
How do you handle feedback when a student has failed?

15. How do you think the new SOL tests will influence your classroom assessments?

16. Use the following scale to answer a few questions about factors that contribute to semester grades you will give to the class you have described. After using the scale provided, then estimate the percentage that factor contributed to the final grade.

1	2	3	4	5	6
Not at all	Very Little	Some	Quite A Bit	Extensively	Completely

2a. **student effort – how much the student tried to learn**

What percentage of the final semester grade is based on student effort?

2b. assessments that measure student recall knowledge (content)

What percentage of the final semester grade is based on assessments that measure recall knowledge?

2c. performances on quizzes

What percentage of the final semester grade is based on performances on quizzes?

2e. objective assessments (e.g., matching, multiple choice, short answer)

What percentage of the final semester grade is based on objective assessments?

(tear off and give to teacher)

1	2	3	4	5	6
Not at all	Very Little	Some	Quite A Bit	Extensively	Completely