MORAL DIMENSIONS OF GRADING IN HIGH SCHOOL ENGLISH

Laurence G. Zoeckler

Submitted to the Faculty of the University Graduate School
In Partial Fulfillment of the Requirements
For the Degree
Doctor of Philosophy
In the School of Education
Indiana University
August 2005

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the
requirements for the degree of Doctor of Philosophy.

_____

Cary A. Buzzelli, Ph.D., chairman

_____

Christine Bennett, Ph.D., second reader

<u>Doctoral Committee</u>:
Prof. Cary A. Buzzelli, Education Dept.
Prof. Christine Bennett, Education Dept.
Prof. Terrence C. Mason, Education Dept.
Prof. Nicholas Williams, English Dept.

_____

Terrence C. Mason, Ph.D., third reader

<u>Date of Oral Examination</u>:
August 18, 2005

_____

Nicholas Williams, Ph.D., fourth reader

Dedicated to my mother, Mary Elizabeth Gallagher Zoeckler, 1921-2001.

Acknowledgments

Thanks to my brothers, Raymond, Philip, and Joseph, and my sisters, Mary Therese, Rita, Elizabeth, and Catherine, who waited so long and so patiently for this work to be finished—or begun, since this is really only a beginning, depending on how one looks at it.

Thanks to Shihong Steve Chen, Ph.D., for all his help unraveling the mysteries of autoformatting; and to those of my colleagues at Utica College who never gave up on the project or my ability to get it done—junior faculty and senior, administration and staff, in the education department and in other departments.

Special thanks to the English teachers of "Mellmax High," whose candor, determination, and high moral standards serve as an example for teachers everywhere.

Special thanks, too, to Mary Drucker, whose constant support and encouragement are deeply appreciated.

And thanks especially to my committee members, Prof. Cary Buzzelli, the ideal committee chair if ever there were one, and Prof. Christine Bennett, Prof. Terrence Mason, and Prof. Nicholas Williams—together, undoubtedly the best such committee ever formed in any discipline at any university.

Laurence G. Zoeckler

MORAL DIMENSIONS OF GRADING IN HIGH SCHOOL ENGLISH

ABSTRACT

Moral issues are deeply embedded in the grading practices of high school English teachers. The problem of giving the right grade for each pupil is central to the study's examination of grading practices of the English department in an upstate New York high school. Arriving at a fair grade, weighing in both achievement and non-achievement factors such as effort and attitude in determining grades, and the role of teachers' expectations in terms of perceived student ability and progress are examined using a theoretical framework derived chiefly from Jackson, Boostrom, and Hansen's *The Moral Life of Schools* (1993). The framework considers the grading process in terms of truth, worthwhileness, trust, and intellectual and moral attentiveness.

A series of semi-structured interviews conducted at regular intervals over the course of an entire school year provide the data for examining the teachers' grading practices and perspectives.

Results indicate that English teachers struggle with issues of fairness, but are confident that their grades communicate the messages they hope to send. Grading strategies are adjusted depending on purpose, and are sometimes altered due to school district grade reporting procedures. Early in the school year, grading is used to help establish expectations. Later in the year, grading is based on the expectations developed from both earlier student performance and personal interaction with students. Grades are subtly influenced by issues of effort, attitude, and conduct, and thus may unconsciously reflect judgments made by the teachers on the moral character of their students. While the teachers acknowledge English class as a proper forum for the exploration of moral issues and the development of character, they hesitate to make direct judgments about the moral development of their students, even as they attempt to influence it.

TABLE OF CONTENTS

CHAPTER ONE: EXPLORING MORAL ISSUES IN GRADING HIGH SCHOOL ENGLISH

Introduction

Grading in a secondary school English class is an activity replete with moral issues. Determining what is right or wrong in any given situation is a complex task when one considers the multitude of issues that may be involved. Consideration for the common good along with that of the individual, the desire to both encourage effort and reward achievement while promoting improvement and further development, accountability to the larger community, the school organization, and parents' wishes, and the satisfaction of one's own conscience, are all part and parcel of the everyday decision-making of teachers. All of these factors are integrated into the process of determining grades.

The purpose of this examination of moral issues surrounding grading in a secondary English program is twofold: first, to explore some of the major issues which inform the day-to-day moral considerations teachers must weigh in their decisions about grading, and, in doing so, to shed light upon the moral complexities of the teacher's role; and second, to give shape to an inquiry into the ways in which teachers address the complex moral issues embedded in their daily routines through the vehicle of grades.

Giving grades is one of many activities associated with teaching. It is commonly taken for granted that grades are merely a kind of mathematical expression of the learners' progress in any given class. For some kinds of courses, like geometry or algebra, the conversion of student performance into grades is accomplished by averaging the scores of tests and quizzes, and the "best" student gets the highest grades because he or she performs the best on the tests and quizzes. The relative weighting of particular

questions, if they are weighted, generally reflects the anticipated level of difficulty, and "right" answers are easily distinguished from "wrong" ones.

In other kinds of courses, however, student performance is not so easily converted into grades. While some tests can be scored much like math tests, so that a given answer will garner a given number of points, many kinds of assignments pose unique difficulties in assigning point values to answers. Term papers, essays, journal assignments, and even "short answers" requiring only a few sentences may have wide ranges for acceptable answers which require subjective responses from the grader, making the assignment of grades less certain. Divergent answers require judgment, and a teacher's feelings about these answers can often affect the assignment of grades to a given answer. A creative writing assignment may receive an "A" from a teacher in one class and a "C" from a teacher in another class. English classes pose a special problem when converting student performance to grades because of the wide variety of assignments which require judgment and subjective as well as objective decision-making.

Assignments which require personal or interpretive responses from the students (and then, in turn, from the teacher) are common in high school English courses. As only one part of an entire school program, however, the performance of students in English must be converted to grades which, on paper at least, look very much like those from a student's math or science classes. Parents, counselors, school administrators, and even teachers and students themselves are prone to "equate" grades across subject areas, guided by an assumption that the processes through which the grades were derived were similar from subject to subject and from teacher to teacher. But anyone who has been to

school knows this assumption is a false one, and grading is a different process from subject to subject and teacher to teacher.

Determining grades for individual students in English classes is an especially complex process, and one which English teachers may work on for many years in order to become "comfortable" with the question, "What's the 'right' grade to give?" At its roots, this is a moral question, and one which is resolved regularly by teachers with every grade they assign. Yet the moral dimensions of grading are so deeply embedded in daily routines and practices that they are rarely consciously examined. This study seeks to explore the implicit moral issues involved in giving grades in high school English, and to shed light upon the ways teachers go about making decisions about which grade is "the right one" to give.

Research Questions

Brookhart's (1991) critique of prevalent grading practices describes the usual method as resulting in a "hodgepodge grade of attitude, effort, and achievement" (p.36). This may be too harsh a characterization, which fails, in the final analysis, to make full sense of the issues carefully weighed out by most teachers in order to come to a grade that reflects, to their satisfaction, a true picture of each student's performance as a whole. In an attempt to explore the "hodgepodge grading," Brookhart (1994) decries, Cross and Frary (1999) suggest that "recommended practice would urge teachers to ignore ability when determining grades" (p. 58). Further on in their examination of typical grading practices, they also recommend that growth, improvement, and effort should be ignored as well when determining grades. In addition, "conduct and attitudes ought to be dealt with separately and not considered when determining grades" (p. 59). This exhortation

demands that a grade become a kind of scientific measurement of specific content—which may not be a suitable way of looking at the range of skills students are expected to develop through their English classes. Cross and Frary's understanding of grading apparently does not correspond with that of many teachers; clearly, English teachers' understanding of what grades mean and how they are meant to be understood is an area in need of exposition, since English teachers cannot ignore the issues of growth, effort, and improvement, and conduct and attitude are arguably significant factors in judging classroom performance.

Thus, the study which provides data for this dissertation aimed at exploring grading practices and issues of judgment, communication, and character development through grading. Specifically, the study began with these central questions:

"How do English teachers decide what is right in giving grades to individual students?"

"How and what do English teachers intend to communicate to students through grades?"

"Is character development a part of the role of an English teacher, and if so, does this influence grading, and how?"

While these questions seem, at first glance, to be so basic and obvious that one could assume that every teacher has grappled with the underlying issues before (or at least, while) grading, the purpose of this study is to make the implicit consideration of the issues explicit in the hope that this will make it possible for the wider academic community, and perhaps even the general public, to understand more clearly the

processes and issues involved in assigning grades—as the teachers themselves understand them—in the context of the English classroom. Feldman, Alibrandi, et al (1998) examined grading practices among high school science teachers in the hopes of providing "an in-depth look at the ways in which teachers use the information they have about their students from various sources . . . to arrive at a summative evaluation of student achievement in class" (p. 141). While Feldman, Alibrandi, et al asked "How do [science] teachers decide what grade to put on students' report cards?" this study hopes to go beyond examining the laundry list of devices teachers use to amass points and calculate averages, and to broaden our understanding of the *meaning* of grades among the high school English teachers participating, exploring both the "how" *and* the "why" of grading.

More importantly, this study examines grading issues in terms of the moral. The grading process is moral because it establishes rankings among students, so grades send messages about the worth of a student in the eyes of the teacher. Because grades are a judgment not only of academic achievement, but also, in a more limited way, of the character of the child who receives them, grading is a moral activity sending moral messages. And because grades can have an effect on the future course of a child's schooling, and to a lesser degree, on the path that child may take in life, grading involves deep moral responsibilities.

Relevant Considerations

Since the turn of the century, grading and giving marks in school has been a source of controversy (Cureton, 1971; Ebel & Frisbie, 1986; Hopkins, Stanley, & Hopkins, 1990). Ebel and Frisbie attributed much of the controversy in grading to three factors:

"the technical challenges of accurately measuring achievement, variations in educational philosophies among teachers, and the conflict in roles teachers' face when they must act as both advocates [for,] and judges [of, their pupils]" (in Cross and Frary, p 53). These three factors are implicit influences in grading in high school English classes; this study seeks to shed light on the ways in which individual teachers balance these factors in order to determine their students' grades.

At a level even deeper than the "educational philosophy" of a teacher, grading is heavily influenced by the values and beliefs of the teacher who grades. A teacher's ideas of right and wrong, of good and bad undoubtedly figure—perhaps significantly—in the evaluation of student work, and in the grades such work receives. In addition, the social aims of the school system can influence a teacher's grading of some assignments. Because grading involves questions of what is fair, what is good, and what helps to form good character in students, it is an activity with deep moral dimensions. While teachers often struggle in making judgments concerning grades (Buzzelli & Johnston, 2002), the moral dimensions of this struggle often go largely unexamined (or at least unsorted), even in times when evaluation and assessment of learning occupy a central position in educational debate. The teachers involved in this study welcomed the opportunity to reflect on their grading practices—and to share their dilemmas in finding a way to accurately convey their assessments of their students' work through the instrument of grades.

Basic Assumptions

While students, parents, and school board members know implicitly that some teachers are "hard graders" and others "easy graders," and (perhaps accurately) ascribe

the differences to personality and outlook differences, a reduction of grading to a simple reflection of personality can make grading appear to be a matter of arbitrary caprice. Such an unexamined explanation reveals little, if any, of the complex decision making and weighing of factors which goes on as a teacher determines what "the right grade" for a given assignment turned in by a given student should be.

While grades can and should be an indicator of a student's mastery of content or the learning of a new skill, and one can argue about what proportion of the fixed-upon grade indicates this part of the classroom teacher's judgment, the grade also reflects many other considerations of differing importance to different teachers. Grades, then, are a composite of several elements, of which mastery and accomplishment is a major part, but not the only part. As an analogy, one might liken it to the way one compares new cars when trying to decide which one to purchase. Yes, the horsepower counts for something, but so do the handling, the safety, the fuel economy, the braking system, passenger room, and yes, the appearance—and even the color. We take the dealer's word for which one is the best knowing that his perspective is not ours, and we weigh that into our understanding.

The supposition that "a grade is a grade" is one that arguably guides public perception about schooling, making it possible to compare and equate grades in different subjects and in different schools with relative ease and considerable self-assurance. Counter to this supposition, this study begins with these basic assumptions:

- grading in an English classroom is a moral activity, involving considerations of "right" and "wrong" on several levels, and at least potentially affecting the formation and development of individual character in the students being graded;

- grading carries consequences for future placement and direction in high school studies, and is seen by many teachers as an influence on the future of students' lives;

- grading can be a very involved process incorporating many finely focused and individualized considerations;

- the interpretation and accurate understanding of a grade requires an understanding both of the student receiving the grade and the teacher giving it;

- achievement is only part of the formula for the grade, albeit a major one. Reducing the grade to this element alone does an injustice to the student/teacher relationship by eliminating other elements that should inform evaluation, whether formative or summative.

Elements of Grades in English

Perhaps more than any other subject, English (or Language Arts) presents complexities in grading. While English has many components which can be graded objectively—spelling words, vocabulary items, comprehension questions based on reading, and so on—it also has many components which are highly subjective and call for judgment on the part of the teacher. Essays, compositions, and even short answers which involve only a few words can be open to interpretation and judgment. Even the more objective components may be incorporated into a larger assignment which is fundamentally subject to judgment: a term paper may lose points for spelling errors, grammatical weaknesses, and incorrect punctuation, and yet be deemed satisfactory in terms of style, expression, or content. An entirely novel composition may be ranked above a technically proficient, but uninspired, piece of prose. How does the English teacher sort all this out? The purpose of this study is to shed light upon the processes

English teachers follow in assigning grades through the complex tangle of elements and interests involved.

In addition to decisions about relative values of individual assignments and quality of work, other factors enter into grading decisions in high school English. Contrary to the admonishments of Cross and Frary, consideration for both the good of the individual and the larger community, the desire to promote achievement and further development while encouraging effort and rewarding achievement, and accountability to the school organization, parents' wishes, the larger community, as well as the satisfaction of one's own conscience are all factors in the everyday decision-making of English teachers, and all of these issues are integrated into the process of determining grades.

Motivation, interest and enthusiasm, participation, consideration for others and positive interaction with classmates can be peripheral to grades in most cases, yet become important issues for certain individual students' grades. This is especially true in the "inclusive classroom," where goals and objectives for individual students' growth and development may be chiefly social, rather than academic. This study examines and compares the ways in which several different teachers grappling with these issues resolve them to their own satisfaction.

Philosophical Underpinnings

Schools have always been charged with the moral development of the children of any organized society, and philosophers since the days of Plato and Socrates have expounded upon the moral aspects of teaching. The question, however, of what constitutes "the moral" is one which seems to have as many answers as there are people to answer it. Definitions of what is moral invariably involve the use of terms which are

themselves open to interpretation, and while some roughly-agreed-upon conception of the moral clearly operates in any discussion of schooling and its purposes, the current debate about what constitutes proper schooling for the young highlights the wide variation in understandings of what is moral or ethical, sometimes giving rise to heated debate and controversy (Beyer, 1997). That books on child rearing have recently been authored by people from camps as widely different as those of William Bennett (1993) and Hilary Clinton (1996) demonstrates that the interest in moral upbringing of children is not restricted to academics, and is subject to ambiguities and widespread disagreement.

If grades serve as the chief means of informing students and their parents about progress in schooling, then it is logical to assume that grades contain some at least implicit dimension which reflects a judgment by the teacher concerning the moral development of his or her students. While academic development is explicit in grades, schools and communities clearly *expect* moral development as well; this is implicit. Yet, perhaps because of its implicit nature, this dimension may not be clear to either the student or the parents—or even to the teacher himself, for that matter. This study attempts to bring this implicit dimension of grading to the surface, and to explore its effect on grading practices, while bringing it to the attention of others beyond the classrooms involved in the study.

Societal concern about the moral development of youth naturally makes the schools an arena for debate over "character education" and issues of morality. While some authors argue that the current culture is one which enforces a "values-neutral" stance upon the schools (Delattre & Russell, 1993) and results in a system where teachers no longer know what their role in moral education is (Gecan & Mulholland-Glaze, 1993),

many others argue that schooling is by its very nature a moral undertaking (Jackson, Boostrom, & Hansen, 1993; Sockett, 1993; Goodlad, Soder, & Sirotnik, 1990). At best, the "values-neutral" stance is a mere posture; any decision involving someone else's good—the basis for every act of teaching—is at its root a moral one, and whether the moral aspect of a decision is explicitly considered or not does not alter its underlying moral features—it simply makes the policy underlying moral decisions unclear.

The embedded nature of moral instruction in the high school curriculum was clearly a distinctive feature of American schools by the time they had become a fixture of the educational landscape. In <u>Moral Principles of Education</u> (1909), John Dewey complained that

> The same distinction between "moral ideas" and "ideas about morality" explains for us a source of continual misunderstanding between teachers in the schools and critics of education outside the schools. The latter look through the school programmes, the school courses of study, and do not find any place set apart for instruction in ethics or for "moral teaching." Then they assert that the schools are doing nothing, or next to nothing, for character training; they become emphatic, even vehement, about the moral deficiencies of public education. The schoolteachers, on the other hand, resent these criticisms as an injustice, and hold not only that they do "teach morals," but that they teach them every moment of the day, five days in the week (p. 3).

This argument that moral instruction is an implicit and underlying feature of all teaching in American schools forms a line of reasoning about the role of schools in the

formation of character that continues unbroken to this day. McClellan (1992) provides an overview of the history of moral education in the United States, beginning with colonial times. He demonstrates that the underlying concepts of democracy were not political so much as moral—equality, fairness, honesty, respect for others, tolerance, and a willingness to compromise where compromise was possible (p. 26-28). In considering the role of teachers in forming the moral character of students, most thinkers avoid the proposal of a course specifically listed as "moral instruction" for use in the public schools. Where such courses do exist, however, they typically focus on the development of character in terms of the moral virtues just listed (Ryan & Bohlin, 1999). Such courses can readily be found in religious schools (Jackson, Boostrom, & Hansen, 1993), and the more specific a religion's teaching on morals is, the more defined the content of such courses can be. Due to the wide disagreements which exist across cultures and religions about what is moral, the public schools, while acknowledging the community's expectations to train students in moral conduct, often attempt to steer clear of direct moral instruction and carry out such instruction indirectly. The bulk of such indirect instruction finds its way primarily into the social studies and English curricula, as the teachers in this study readily confirm.

This indirect approach to moral education requires an establishment of some basic notions about the moral, and central to most definitions of the moral are concepts of right and wrong and considerations of what constitutes desirable ends and the means to achieve those ends. In discussing the moral basis of teaching, Tom (1984) explains the moral as "a concern for the rightness of conduct and a broader concern for what is deemed important or valuable, provided that these valuational situations clearly entail

desirable ends" (p.70). Drawing on Rawls (1971), Kerr (1987) asserts that "A well-ordered society both advances the good of its members and is regulated by a public conception of justice" (p. 23). "Education," according to Kerr, "is an initiation into a culture" (p. 23). The moral features of that culture must be passed on, and in doing so the teacher's obligation is both to the students and to the culture of which they are a part. But another aspect of moral training has to do with learning how to treat others in a way that corresponds with the deeper values of a society; teachers must model such treatment, and make their students cognizant of their duty to treat one another in ways that are "right." Nel Noddings (1984) considers these issues in her reflections on grading in terms of a caring relation (pp. 193-196). Her maxim that the "student is infinitely more important than the subject" (p. 20) suggests an element in grading that adds significantly to its complexity—how do teachers grade in a way that ultimately benefits the student as a person while staying within the limitations imposed by the school reporting system?

Clearly schools depend on a link between character and caring for others, evident in the fact that parents entrust their children's formation, in at least some part, to their local schools. Noddings (1995) carries this link beyond the traditional disciplines of education, and argues that schooling should be organized around themes of care. "All children must learn to care for other human beings, and all must find an ultimate concern in some center of care" (p. 366). Her vision "in favor of greater respect for a wonderful range of human capacities now largely ignored in schools" (p. 366) rests on the belief that the fundamental purpose of schooling is a moral one, and that "skills education" and "calls for excellence" in education fail to address the ultimate purpose of such efforts. With current movements in education toward "higher standards," English teachers must

balance their traditional interests in promoting human sympathy through literature with boosting the level of student performance in basic skills; how that balance is maintained or imperiled in the present is another element of exploration to be examined in this study.

Baier (1995), explains trust as "letting other persons . . . take care of something the truster cares about, where such 'caring for' involves the exercise of discretionary powers" (p. 105). The discretionary powers of teachers in giving grades—especially in the English classroom—are broad, and this study will explore the range of discretion that the participating teachers believe is appropriate and acceptable to their own consciences in their grading practices.

Taking these implicit features of grading into account, the moral weight of the process becomes significant. As Buzzelli and Johnston (2002) put it, "the act of grading, then, is a moral one *par excellence*" (p. 60). The extent to which the high school English teachers involved in this study are cognizant of a personal obligation toward their students and a wider obligation to society will be explored and delineated.

Theoretical Framework for the Study

The Moral Life of Schools

The theoretical framework which forms the foundation for this study is drawn chiefly from the work of Philip Jackson, Robert Boostrom, and David Hansen, whose book, *The Moral Life of Schools* (1993) examines moral aspects of schooling, observing that the effects of schools extend far beyond explicit attempts to assert a moral influence on students. Without consciously intending to act as moral agents, teachers exert a moral influence through all that they say and do in their interaction with students. The framework developed by Jackson et al to explore moral issues in instruction is extended

in this dissertation to facilitate the exploration of moral issues in grading (specifically, in high school English), an issue left almost entirely unexamined in *The Moral Life of Schools.*

Although Jackson, Boostrom, and Hansen did not specifically address grading as a moral activity, they developed a lens through which most school activities, including grading, can be examined for their moral implications. The Moral Life of Schools Project, the landmark study which formed the basis of Jackson et al's text, studied several schools and many teachers and classrooms. The study yielded several "categories" of school activities which the researchers identified as explicitly moral. Jackson et al grouped the categories under two larger headings which they labeled "moral instruction" and "moral practice."

Moral instruction included these categories:

- "moral instruction as a formal part of the curriculum;"

- "moral instruction in the regular classroom;"

- "rituals and ceremonies;"

- "visual displays with moral content;" and

- "spontaneous interjection of moral commentary into ongoing activity" (p. 42).

The activities falling within these categories generally constitute part of the larger school agenda for promoting moral development in the students. Teaching students right from wrong, honest from dishonest, and fair from unfair is part and parcel of a complex web of discourse woven into the day-to-day activities of school. Many of these activities can be seen as being tied at least indirectly to grading or giving grades. Even apart from

grading, they serve to establish quite clearly the moral nature of schooling and the pervasive nature of moral undercurrents in the day-to-day activities of schools. Jackson et al. then introduce three final categories which constitute "moral practice." It is here that the moral dimensions of grading become especially evident. Moral practice encompasses the following categories:

- "classroom rules and regulations;"

- "the morality of curricular substructure;"

- "expressive morality within the classroom" (p. 42).

Complexity of Schooling

The moral practice categories, "rules and regulations", "the morality of the curricular substructure", and "expressive morality within the classroom" figure significantly in the day to day activities of classroom teachers, and while Jackson et al. do not examine grading in detail using these categories, this dissertation makes grading its focus, thus building upon the foundation established in *The Moral Life of Schools* to expose the moral issues involved in giving grades and using them to foster the development of students according to the better judgment of their teachers.

Rules and Regulations

In "Classroom rules and regulations" (pp. 12-13), Jackson et al observe that "Every classroom constitutes a small society embedded within a complex web of social entities whose overlapping systems of laws, customs, and traditions it partially shares and sometimes adds to or contradicts" (p. 12).

While Jackson, Boostrom, and Hansen are discussing rules of behavior and procedure (raising hands before speaking, fire drill rules, etc.), the observation is just as pertinent to grading and grading practices: every teacher creates a system for grading which overlaps with the systems of the other (English) teachers, whether of the same grade level or across grade levels; with the school's grade reporting system; and with the grading systems used by other teachers in the other subject areas. At the same time, however, the individual teacher creates a unique system which may add on to, or take away from, other teachers' systems, and it is common to hear students inquire first, "What grade did you get in English?" and next, "Who's your teacher?" indicating that the grade is only fully appreciated or understood when the giver is known.

Although Jackson et al are discussing the application of classroom behavior rules, their remarks concerning the complexity of such application again transfer readily to grading:

> "…as simple and as direct as rules sound when they are put into words, . . . they turn out to be quite complicated when we try to understand their enactment. This is partly because most rules seem, at first, to be inconsistently enforced. . . .What gradually becomes evident, however, is that many of these apparent inconsistencies are not actually instances of rules being ignored. Instead they reflect refinements of the rules that are clearly understood by almost everyone present except the observer. In other words, the rule as stated turns out to be a general maxim to which there are many exceptions" (p. 13).

A similar observation can be made of grading; the application of the rules may seem inconsistent to the outsider, yet few students dispute the grades that appear on their report cards because each understands how the grade was arrived at, or at least believes that the teacher applied the rules correctly in accordance with a shared understanding of how the grades were to be derived. Where this does not take place, disputes over grades and the fairness of the procedures occur.

Curricular Substructure

An examination of the "morality of the curricular substructure" (pp. 14-29) provides an important component of the framework used here. Unlike the previous categories, which deal with essentially overt efforts to influence students, this final category introduces elements of schooling that highlight the complexity of the moral influences that affect grading practices.

Jackson et al introduce the topic of the morality of the curriculum substructure by observing that the curricular "content. . . is not all there is to it, for in addition to containing content, every curriculum is also structured in a variety of ways" (p. 14). Considering the more obvious ways for organizing content, the authors identify principles of organization such as chronology, narrative, and topic. Movement along a time line or from one discrete topic to another reveals this kind of organization; some subjects, like math, are ordered "according to level of difficulty . . . first are the easy problems, then the hard ones" (p. 14). While the organizational structure is sometimes very obvious, as in history where timelines are important, and "the structure. . . becomes part of the lesson" (p. 14), in other classes, the structure is less visible, especially where the content is arranged topically, and teachers depend on a textbook to provide the organizational

backbone of the curriculum.   Here, "the structural elements of the curriculum that determine the order in which the material is taught receive scant attention during the lesson itself and thus remain barely noticeable to those present, including, it would seem, the teachers themselves in many classrooms" (p. 15).

Typically, high school English curricula are a hybrid of organizational schema: broad topics—often identified as "units"(short stories, drama, novels, poetry)--form an umbrella over many elements that require ongoing development of skills (reading, writing, spelling, listening, speaking) which themselves may require cognitive exertions ranging across Bloom's entire taxonomy (knowledge, comprehension, application, analysis, synthesis, and evaluation). All of these call for assessment on the part of the English teacher, and the context of schooling demands that such assessment take the form of grades. Thus, the principles of organization governing the high school English curriculum both shape grading practices and are shaped by them in varying degrees from assignment to assignment and from unit to unit. A basic element in the framework guiding this study, then, is an examination of the outward structures of the teachers' grading systems and their relationship to the formal content presented in the teachers' classes. The ways in which these structures influence the grades awarded—or earned—brings some of the moral issues tied to grading to the surface for discussion. Even the choice of verb—"awarded" or "earned"—indicates a moral leaning, suggesting a carefully nuanced understanding of responsibility for the outcome, in that it is the *teacher* who "awards" a grade, but the *student* who "earns" one.

This fundamental element of grading—the outward structure—is often as much as most "outsiders"—parents, school administrators, other classroom teachers, and even

many researchers—believe they need to know in order to understand how grades are determined and to judge whether or not such grades are fair—essentially, the crux of the matter when determining the moral soundness of grades and grading practices. But this is only the first step, as Jackson, Boostrom, and Hansen reveal as they delve more deeply into the morality of the curriculum substructure.

Jackson et al. label their "shaping forces" substructural, in part to distinguish them from the outward organizational principles governing curriculum arrangement ("structure") and in part because these substructural forces seem to lie beneath or behind the outward organization:

"What prompted us to think of them as structural was the way they seemed to lend a kind of helping hand to the ongoing activity, buoying it up like a cushion of air or a buried foundation of some kind, a layer of bedrock, perhaps. In fact, they often seemed to reside so far beneath the surface of what was being done and talked about within the lesson that we ultimately abandoned the term *structural* in favor of *substructural*" (p. 15).

Remarking on the "near invisibility" of these elements, Jackson et al explain that "they are seldom explicitly acknowledged by either teachers or students except when they are absent to start with or when things go wrong" (p. 16). Arguing that these features are essential if instruction is to take place, they are further elaborated as

> "part of a complex web of obligations and responsibilities whose strands interlock to form a kind of moral substratum of understanding . . .The most helpful approach we have found so far in trying to understand these enabling conditions is to envision them as composing an elaborate amalgam of shared understandings, beliefs,

assumptions, and presuppositions, all of which enable the participants to interact amicably with each other and work together" (p. 16).

An extended exploration of this substructure reveals "shaping forces that . . . remained out of sight much of the time yet continued to function behind the scenes and that appeared to be crucial to a full understanding of what was going on" (p.15). Again, Jackson et al are concerned with instruction rather than grading, yet what they discovered can be used to shed light on the moral complexity of grading, especially in high school English. The role which the embedded "shaping forces" play in instruction is a major one, according to Jackson et al, yet the role these forces play in grading and in the *interplay between grading and instruction* serves, in fact, to bolster their point that the moral life of schools is indeed "much more complicated" than an explanation of organizational principles alone can reveal. *The Moral Life of Schools* focuses on instruction and teacher-pupil interaction to reveal forces shaping moral instruction; this dissertation incorporates those forces into a framework which allows an examination of the moral issues of grading by examining the ways in which those forces both shape grading and are strengthened—or weakened—by grading practices.

Truthfulness

The first of the substructural elements or enabling conditions presented is truth-telling. "Teachers the world over are expected to speak the truth when addressing their students, and students are expected to do the same when speaking up in class" (pp. 16-17). The discussion which follows supports the argument that truth-telling is an expectation that is essential if instruction is to take place. If teachers intentionally mislead or misinform their students, they are clearly morally deficient. If students do not

tell the truth about what they know or do not know, instruction is subverted. However, the authors contend, "most of the time, … teachers and students alike can only presume that the condition of truthfulness is being upheld" (p. 17). Rooting out lies and liars, the authors observe, can require a lot of time, will disrupt instruction, and, if handled poorly, "can irreparably damage the quality of the interpersonal relationship between the two parties" (p. 17). Instances where truthfulness is called into question must be kept to a minimum if instruction is to proceed smoothly, and the "*assumption of truthfulness*" is posited as a basic moral element of instruction. Jackson et al argue that this same assumption guides most social interactions, but that different contexts alter the moral coloration of the assumption: doctors and patients, for example, interact on the assumption of nearly total truth-telling; diplomats and used-car buyers, on the other hand, interact with an assumption of only partial truth-telling (p. 18).

While Jackson et al. concede that the assumption of truthfulness cannot always be supported in the classroom, they demonstrate that when it cannot, instruction falters. Without discussing issues of grading per se, the authors do remark on testing as one of the times when teachers feel compelled to enforce the assumption of truthfulness, often by reluctantly revealing less than complete trust in their students' truthfulness or trustworthiness. Naturally, the assumption of truthfulness can be applied to grading practices and grades themselves.

Worthwhileness

The second substructural element introduced in The Moral Life of Schools is the "*assumption of worthwhileness*," summed up by Jackson et al as "the mutually shared assumption that the material being taught is important and the activity being engaged in is

worthwhile" (p. 24). Again, the authors posit that the absence of worthwhileness "practically ensures instructional breakdowns and difficulties of one kind or another; it is difficult to imagine how either teaching or learning could occur. . . if either teacher or students totally lacked the conviction that what they were doing was worthwhile" (p. 24). To further support the claim that worthwhileness is a morally charged concept, the authors argue that

"What makes the assumption of worthwhileness moral is the even more deeply embedded assumption on which it rests, which is that schools and classrooms are places where one goes to receive help, to be made more knowledgeable and more skillful. Schools and classrooms are designed to be beneficial settings. This implies that the people in charge care about the welfare of those they serve and only ask them to do things that are expected to do them good. Without that underlying assumption, schools start to resemble prisons, which is how they must begin to feel to those who lack faith in the institution's good intentions" (p. 25).

The assumption of worthwhileness is "a tacit acknowledgment of the moral character of the institution" (p. 26). While most students, parents, and community members are convinced of the worthwhileness of school—or merely take it for granted— where they are not, grading becomes suspect and poor grades often reflect not so much a failure to achieve on the part of a student, but a failure to "buy into" the worthwhileness of schooling.

Trust

Trust is a third substructural element guiding moral interaction in schools. Clearly, both the assumption of truthfulness and the assumption of worthwhileness depend on an

operative relationship of trust between teachers and students, so that trust can be seen as an important substructural element, whether as a prerequisite for the other two, or as something that grows out of them. While Jackson, Boostrom, and Hansen mention trust in relation to these assumptions, they do not treat it as fully. Nevertheless, a brief consideration of almost any instructional situation calling for either truthfulness or worthwhileness will make it clear that trust is required (even if in differing proportions) between the teacher and the taught. The work of Noddings (1984, 1995) underlines the importance of developing a relationship of care between teachers and students; clearly, trust is a part of the foundation of such a relationship.

Trust removes a large part of the burden of determining the truthfulness and worthwhileness of any given piece of the larger picture. If students can trust their teachers, and vice versa, and parents can trust their schools, then many minor controversies which might arise from second-guessing what is being done in school can be avoided. Trust is clearly a prerequisite for grading, and the ways in which the participating teachers develop the trust needed to make grades meaningful are examined in this study.

Expressive Morality within the Classroom

Jackson et al. include a final category under their heading of "moral practice in the classroom." Beginning with a discussion of how teacher's express some forms of judgment and reveal their thinking through facial expressions, they develop the notion that certain kinds of messages are sent through "Looks of kindness, impatience, good humor, sternness, incredulity, indignation, pity, discouragement, disapproval, delight, admiration, suspicion, disbelief—the list could easily go on" (p. 30). The teacher's facial

expressions, combined with gestures and postures, constitute an "expressive morality." Applying the notion of expressive morality to the classroom teacher, Jackson et al observe:

"As an instance of how broadly the net of observation might be cast in search of expressive meaning, consider the well-known sense of trust (or distrust) in a teacher that students and outside observers gradually develop. Here, surely, is a morally relevant quality if ever there was one. And for no one more than a teacher. After all, what greater virtue than trustworthiness could we ask of a person who is in a position to so potently influence the young?" (p. 33).

Trust is certainly something that teachers convey to students in expressive ways, and trust between students and teachers is especially important in the matter of assigning grades. A remark like, "please do your own work," written across the top of an essay sends an important message of distrust; a teacher whose grading seems subject to caprice or driven by preconceptions cannot hope to gain the trust of her students. At the same time, the students must gain the trust of their teacher, and grades can sometimes hinge on how well this trust is maintained.

The embeddedness of moral issues in schooling is not merely a matter of philosophical or theoretical interest; if the underlying moral perspectives of a school system can be determined, they may help to explain why some courses of action are pursued while others are not. District grading policies, both stated and implied, will be explored in order to determine their impact on grading practices.

Applying the notions of truthfulness, worthwhileness, and trust to instruction can help in judging the moral appropriateness of lessons in general, but considering particular

lessons—and how to grade them—may require consideration of several other factors, such as state mandates, school imposed curricula, testing requirements, and so on. All of these considerations come into play in determining grades; sometimes one feature (say, trust) must carry the weight of the decision to grade an assignment despite the suspicion that another feature (say, worthwhileness) is weak—hence, a requirement driven by a state testing mandate may not seem worthwhile to the either the students or the teacher, but the students' trust in the teacher's grading practice will be enough for them to take the preparatory assignment seriously.

Additional features of the philosophical framework

For purposes of this examination of moral issues in giving grades, the three elements embedded in the philosophical roots of schooling serve as a basic framework for examining the moral dimensions of decisions guiding grading practices. This basic framework is bolstered by the addition of Hansen's (1997) concepts of "moral attentiveness" and "intellectual attentiveness." Hansen presents these concepts in a discussion of what it means to understand students. Addressing the question, "What kind of relationship should I form with my students?" (p. 1), Hansen observes that teacher candidates "worry about liking and respecting students—and, in turn, being liked and respected by them" (p. 1). This concern is one of most teachers, not merely pre-service ones, and it can influence grading practices in ways that can have serious moral implications. For example, using grades as a means of winning the affection of one's students without worrying about their validity is certainly wrong; but showing some flexibility in grading—an element of "mercy," as it were—may be the most appropriate—and moral—action in some circumstances.

Hansen aims at establishing a way of looking at teaching which springs from the assertion that "at teaching's core is the idea of serving students' intellectual and moral growth" (p. 1). He argues that "at the center of understanding students is learning how to be intellectually and morally attentive as a teacher" (p. 2).

Intellectual Attentiveness

A brief consideration of the purposes of teaching leads Hansen to assert, quite logically, that teaching of its nature requires "intellectual attentiveness" both to students' responses to the subject matter that the teacher presents to them and to "the persons students are becoming" (p. 4). This is both intellectual and moral, because on the one hand it "presumes the teacher's familiarity with the subject and its logic and structure" (p. 4), and on the other hand, "it entails care and concern for the students in their relation to the subject" (p. 4).

Intellectual attentiveness demands that a teacher pay close attention to students' intellectual processes as they take up what is being taught. It means testing for comprehension, probing depth of understanding, and "being alert to aspects of student conduct that influence their engagement with subject matter" (p. 4). It is thus closely connected with subject matter and teacher expertise in that subject matter.

Moral attentiveness

"Moral attentiveness," the other key concept introduced in Hansen's paper, can be treated separately from intellectual attentiveness "only for heuristic purposes" (p. 10), since moral development accompanies the intellectual growth of the students. However, its focus is slightly different:

"Moral attentiveness has two components: alertness to the development of students' character, and awareness of one's regard and treatment of students—or, put differently, awareness of one's own character as a teacher" (p. 8).

This suggests that grading practices have an affective component—one so deeply embedded as to be largely invisible to a non-reflective teacher, but likely to be detected by a student who senses a hidden bias for or against him or her on the part of the teacher.

The strength of Hansen's concepts resides in the effort to return the focus of teaching to the personal engagement required between teacher and students, and the re-assertion of the teacher's moral obligation to take a genuine personal, yet professional, interest in both the processes of the classroom and the students as persons. Applied to high school English, these concepts suggest that "moral" grading reflects, to a greater or lesser degree, the level of understanding which develops as a teacher seeks to be both intellectually and morally attentive to his or her students as *individuals*.

Taken together, intellectual attentiveness and moral attentiveness can be used as guides in making decisions about instructional techniques. These must be tailored to fit the particular students being taught; sensitivity to both their intellectual development and the way they should be treated are essential to morally sound methods of instruction. Whether and how intellectual and moral attentiveness affect the grading considerations of the participants are explored in this study.

Balancing Multiple Concerns in Grading

The classroom teacher's day-to-day routines of instruction, grading, and discipline are clearly morally laden activities. These activities are so central to the daily operations of

schooling that they are taken for granted. Decisions concerning what to teach, how to teach it, how to keep students involved without disruption, and how to evaluate all of this activity are made every day in every classroom, and undoubtedly exert an influence on grading practices. The framework established above, with its elements of trust, truthfulness, worthwhileness, and intellectual and moral attentiveness, allows for an examination of teacher practices in grading to see how the teachers participating in this study deal with the moral implications of grading, whether consciously or unconsciously. Balancing the multiple concerns involved in grading and still being confident that the grades being given are the "right" ones is a complex, dynamic process.

Implications of grades

There are many aspects of schooling that influence the moral decisions made in giving grades. Because grades are seen as measures of merit, "many youths feel marginal to the central school population partly because they are receiving messages (in the form of failing grades) that they do not belong in school" (Sinclair & Ghory, 1987). School grades may reflect a student's relative performance in the school (Wood, 1994), or may constitute "an easy lie" which tells nothing about their actual performance (Tomlinson, 1994). Grade depression (Wood, 1994) or grade inflation (Bracey, 1994) may creep into school practices as parents demand improved performance and grades seem the only way to demonstrate it. Dockery (1995) points out that grading scales are often arbitrary and vary from teacher to teacher; grades may be used to influence behavior, and thus not accurately depict academic performance, and "zeros are motivation killers" (p. 34). None of these issues is alien to an English teacher, and how the subject population deals with them is examined in terms of the way one kind of decision or another falls within (or

outside of) the framework being used to consider the moral implications of the grading process.

Questions of how to assign relative weights to homework assignments, projects done collaboratively, and scores on tests and quizzes must be resolved. With so much to consider, Hendrickson & Gable (1997) lament that "the exact relationship between student classroom achievement and teacher grading practices is unclear" (p. 159). If this is so, questions of equity, accuracy, and even honesty in grading arise—after all, what are the teachers really doing? While "classroom achievement" seems on the surface to be a simple and easily assessed construct, it is considerably more complex than its label suggests. Is it measured against an expectation of ability, other students' performance, or a student's previous performance? Is it determined against an external standard, or is it based on a shared understanding? Or is it a combination of some or all of these? Clearly, grading is heavily laden with moral considerations which go well beyond the requirements of "achievement testing." This study does not seek to establish Hendricikson & Gable's "exact relationship between…achievement and grading." Rather, it seeks to shed light upon the ways in which teachers grapple with that relationship and many others, attempting to advance their students intellectually, morally, and personally, using grades as an expression of perceived advancement or the failure to advance.

Good student, bad child; Poor student, good child

Discipline is another aspect of teaching that involves moral issues and may creep into grading. While district policies and personal styles of classroom management may influence discipline decisions, they are inherently moral in that they involve decisions

about right and wrong conduct—initially, on the part of the student, but then, in response to this, on the part of the teacher and the system of which the teacher is a part. How discipline issues influence grading issues is also explored, especially since "bad behavior" may be punished through reduced grades, and "outstanding behavior" may be rewarded with extra points. Decisions about discipline are not necessarily self-contained, and teachers must consider likely ripple effects on the other students, other classes, the administration, and the parents. Any of these could rebound on the teacher as well, so prudence in action is essential; if grading is influenced—either directly or indirectly—by discipline issues, then these issues must also be explored.

The complex interconnectedness of the many elements of schooling ultimately finds some expression in the grades which appear on each student's report card. This requires careful weighing of both the implicit and explicit messages that grades can send, consideration of the consequences of giving certain grades to particular students at specific times, and the distillation of many factors into what is ultimately reduced to a single letter or number grade on a report card. This is a case study which examines the whole picture as the teachers who do the grading see it, rather than restricting its scope to an examination of some limited facet or facets of grading. The framework used here explicitly acknowledges the complexity of the moral issues in grading, and this study provides valuable insight into the participant teachers' demanding moral task of assigning grades on a regular, even daily, basis.

CHAPTER TWO: DESCRIPTION OF STUDY

The study which provides data for this dissertation was a case study involving the English department (grades 7-12) at a rural high school district in upstate New York. The school, the English department, and the English curriculum at the school are typical of schools of the same size in the region covering an area of several rural or semi-rural upstate New York counties within a hundred-mile radius. In terms of per pupil expenditures, teacher/pupil ratios, curriculum sequencing, daily school schedule and school year calendar, the district is typical of many in the region.

Mellmax High School

The study took place in a medium-sized (1800 students, grades 7-12) semi-rural public junior/senior high school in upstate New York. Mellmax High School (a pseudonym) is situated in the largest town in the county, and receives a high proportion of state aid in relation to its overall budget (90%+) due to the economically depressed circumstances common throughout the region. The student body is comprised almost exclusively of white students, with a very small minority enrollment (less than .5 %). This is in keeping with the demographic profile of the town and the county itself.

The school is the largest in the county and enjoys a relatively good reputation among the local population, both in terms of academic achievement and the general discipline of the student body. The general sense among the faculty is that the school is on a par with any of its neighbors in terms of curriculum and instruction. The proportion of students receiving free or reduced-cost lunch exceeds sixty percent, and so the school district is classified by the state as a "high-needs district."

The school consists of a single building with a junior high wing and a high school wing. The junior high is made up of the seventh and eighth grades, and the high school of grades nine through twelve.

The English Department

The departmental structure of Mellmax High School is typical of neighboring upstate New York schools, with each discipline having its own department chaired by one of its members, usually a senior member with twenty or more years' experience in the district. This was the case in this study, the English chairperson having served in the district for well over twenty years.

The English department members were loosely allied with one another and supportive of one another's efforts. If teachers had similar assignments, they typically conferred with one another about curriculum and often traded materials when they taught the same units. The chair's chief responsibility was to insure that paperwork (attendance reports, five-week letters, and report cards) was submitted in a timely fashion; day-to-day decisions about curriculum or management issues were left to the individual teachers, who had considerable flexibility in matters concerning their own classrooms. The teachers involved ranged in age from their mid-twenties to early sixties, with years of experience ranging from two years to over thirty years. The group constitutes a purposive sample (Fraenkel & Wallen, 2000) which includes the entire department. All the teachers were certified to teach English in accordance with state regulations.

The Mellmax High School English department consisted of ten teachers; the sample used for this study included all ten, plus two recently retired teachers and one who had recently changed jobs and begun working at a new school. These last three were

included in the sample because their different status provided opportunities for additional comparison. Since the department can be seen as typical of those in many schools in upstate New York, this study can be seen as a kind of case study (Merriam, 1988; Yin, 1989; Stake, 1994).

The retirees provided insight into the long-term aspects of grading policies at the school being studied. Having worked in the district for many years, they could recall how grades had been given in the past and confirm the continuity of the grading process in terms of district rules and community expectations. As a rule, schools change slowly, and teachers with long experience in the same district can provide insight into changes in school policy and student performance over a span of many years.

The job-changer provided information on his grading practices while employed in the subject district before changing jobs. He carried his grading practices to his new district in another county some sixty miles to the west. This allowed for comparison with practices in another school in a different district, and provides limited evidence that grading practices are likely to be more alike than different in districts other than the subject district. Had the new position required notable changes in the teacher's grading methods, it would suggest that Mellmax High School may not be typical in such matters; the fact that the new district took it for granted that the job-changer's grading was appropriate suggests that his methods for grading fall within an acceptable, if not clearly defined, set of such practices. The inclusion of these former department members, then, allowed for limited triangulation of the data (Denzin & Lincoln, 1994) by time (past practice) and location (a different school district).

Data collection involved an extensive interviewing process described in detail below. Each teacher was interviewed several times over the course of an entire school year (2001-2002). All interviews were recorded with the consent of the teachers, and the contents were transcribed for analysis.

The use of the entire English department in the sample supports its purposive nature. This group of teachers presents a variety of views on the issues discussed during the study and is fully representative of the department, as the sample and the department are one and the same. The department, in turn, is similar to other English departments across the state in personnel and structure, approaches to instruction and reading lists. The group's similarity to English departments throughout the upstate New York region makes it a typical case sample.

While qualitative studies such as this are not generalizeable in the way that quantitative studies are, the issues and dilemmas involved in grading are similar from school to school, and insofar as the views of these teachers resonate with those of English teachers in other schools, the case study presented here can offer valuable insight to others hoping to understand how English teachers come to determine the "right" grades for their students. More importantly, this study allows those outside the arena of the English classroom to begin to develop an understanding of the moral complexities of grading faced on a daily basis by all teachers.

My Relationship to the Department

The teachers involved in the study had been my colleagues when I worked in the sample school's English department. I had taught English in the high school for ten years before returning to the university to pursue a doctoral degree. Grading was (and is) a

perennial topic of discussion among the teachers, and when the topic was proposed for a formal study, virtually everyone in the department embraced the idea enthusiastically. My own experience in English grading—working first in the junior high for five years, and then in the senior high for the next eight—provided a sense of the issues involved and helped inform the interview process. I was well versed in the entire English curriculum, and had considerable experience in grading the New York State Regents examinations in English (first for the eighth graders, then for the eleventh—these years being the ones focused upon by the "standards makers").

My personal acquaintance with the parties involved provided both an increased likelihood of obtaining frank answers to all questions (Briggs, 1986) and a risk of bias in interpreting those answers (Bogdan & Biklin, 1998 [1982]). Bias was controlled for by carefully comparing the data across participants and member-checking to be sure that what was recorded in interviews was complete and that summaries were accurate. All participants were given the opportunity to review the data from their individual interviews in order to insure accurate representation of their views.

For purposes of analysis and cross-comparison, the teachers were initially grouped by years of experience. This yielded three categories: "novice," with less than seven years experience; "established," with seven to fifteen years experience; and "veteran," with more than fifteen years experience. While the initial expectations of the researcher were that the teachers within each of these categories would have similar views on the grading issues discussed in the interviews, this proved to be untrue: in each of the experience categories, variations in outlook and practice were found, and the analysis shifted to the collection of data which represented shared notions of practice and purpose in grading

regardless of years of experience. Nevertheless, for purposes of acquainting the reader with the participants, the categories of novice, established, and veteran are used by way of introduction, and are used throughout the study for comparison purposes.

Pseudonyms were assigned to the subjects in order to maintain confidentiality; in order to ease the reader's burden of remembering the entire set of thirteen teachers, the assigned first and last names are alliterative (i.e., begin with the same letter), and experience is indicated by alphabetical position from "A" to "M", that is, the least experienced teachers' names begin with letters at the beginning of the alphabet ("Alice Andrews," Betty Browne," and so on) while those with the greater experience have pseudonyms beginning with letters occurring later in the alphabet, with "Mary Minton," at thirty-three years, having the most experience. "Frank Fender" and "Fiona Fisher" have equal experience at eight years, and fall into the middle of the "experienced" classification. The subjects are introduced below in order of years of experience, beginning with the least experienced. Although New York currently requires a master's degree for certification, this was not the case when the veterans were hired. However, most had earned master's degrees in reading or English literature. All the established teachers had master's degrees, and the novice teachers were either working toward the required master's or had already earned one. Levels of education are indicated in the table below.

Novice Teachers

The novice teachers had less than seven years' experience. All were still trying new methods of grading and sometimes radically altering their approaches to grading, based on their increasing level of experience.

Alice Andrews had taught for only one and a half years. She had previously served in the district as a long-term substitute teacher for three months with 10th grade honors, journalism, and twelfth grade. At the time of the interviews, she was teaching 12th grade. Her grading system was "in transition," i.e., it was being adjusted as she gained experience. "I'm still figuring out what works best—you know, what keeps them working and makes sense."

Betty Browne had taught seventh and eighth grade English in the district for two and a half years. Like Alice, Betty's grading practices were in flux and changed as she learned how different schemes would affect outcomes. For example, her homework policy had been adjusted due to harsh reality: at first, she imposed time limits and deducted points for late assignments, with increasing penalties for increased lateness. These time limits and penalties were imposed until "practically everybody was failing," at which point, she adjusted the policy to prevent such a dire outcome. Of all the teachers in the sample, Betty showed the greatest willingness to experiment with grading schemes.

Catherine Carney had been teaching for three years, and taught classes of ninth and twelfth grade English and an eleven/twelfth grade elective on film. She had begun her career in another district, teaching only ninth grade English. Her "rough balance" of assignments was "about 60% writing, 40% quizzes and handouts, and homework. If you don't count homework, they won't even bother."

David Dutcher taught ninth and eleventh grade English and twelfth grade public speaking. He had taught for five years. "I especially feel comfortable with the 11th grade because it's the one thing that's been consistent throughout the five years." David's approach to grading was quite technical in comparison to many of his colleagues; he

awarded points for component pieces of essays, using a grading sheet and adding the scores on each part to arrive at an overall grade. David was the job-changer, having switched school districts after four years at Mellmax High School. He continued the grading practices he had developed at Mellmax High at his new school.

Ellen Enders had taught for five and a half years, starting her career in another district and coming to the sample district after a year and a half in her initial district. Her change from one district to another had not caused any difficulties in adjusting her grading practices, again suggesting that an acceptable practice for arriving at grades was widely shared across the local school districts. She taught both tenth and eleventh grade, but "I like 11th better. . . it's not the kids, it's the curriculum." For Ellen, grading was in part a function of length of assignment: "Short homework assignments count once; tests generally count twice, and my essays and things that require more of their time count three times. So I weigh them depending on how long, generally, and how difficult the assignment is."

<center>**Established Teachers**</center>

The second category of teachers had experience of more than seven but fewer than fifteen years in teaching English. All expressed some level of comfort with their by now "established" grading practices; all were more assured about their systems than the novices, but less certain than the veterans.

Fiona Fisher At the time of the interviews, Fiona had a total of eight years' experience in teaching English. Her early experience was in the middle school (grades seven and eight), and she had spent her first five years of teaching in another local school district. She had been teaching at Mellmax High School for three years. Here, her

assignments included tenth and eleventh grade English. Fiona had enjoyed teaching in the middle school, but was also happy with the high school position. Her basic philosophy about grades was established in the middle school, and again, this strengthens the observation that general ideas about grading are widely shared among school districts—and even across grade levels. At the beginning of the school year and at each report card, Fiona tells her students that they *earn* their grades. "I personally don't feel that I am giving them a grade. There are so many assignments and they are worth so much. What the students earn for each assignment adds up to their grade in the end. They say to me, 'What grade did you give me?' and I say, 'I don't *give* you grades, you *earn* them.'"

Frank Fender had taught English for eight years, seven of which had been devoted exclusively to twelfth grade "regular" English (i.e., not Advanced Placement). At the time of the interviews, Frank was teaching grades ten, eleven, and twelve, and was ". . . getting comfortable with that." Frank was adjusting his grading practices according to grade levels; for example, "Homework is checks and zeroes; missing homework can hurt their grades. For the 10th graders, homework is part of the package, for the others, it depends on whether they're 'on the bubble.'"

Gail Goodwin had taught eleventh and twelfth grade English for a year before changing schools and assignments. She then taught seventh and eighth grade English for eleven years, and was teaching these grades during the time of the study. She used a "point system" for grading. "Every assignment is worth a certain number of points. Their job is to earn as many, if not all, of the points that I give out. So then I just divide their points by my points, and that's their grade. And I do a sample to show them—to see it on paper. It's in my objectives."

Veteran Teachers

The last category, the veteran teachers, consisted of career teachers who had spent fifteen years or more in the English classroom. Of the five veteran teachers who participated, two were recent (within the previous two years) retirees, and were included in the study in part because of their long experience, and in part because of the belief that, being no longer involved in the internal power relationships of the school, they would be able to speak frankly about school policies and actual practices. They would also be able to provide information on changes in grading practices over the long span of their service.

All the veteran teachers were comfortable with the grade level they had taught longest, and felt well established with regard to course content and expectations. All had taught in the sample district for at least fifteen years. All were female, but five years previous, all the longest-term veteran teachers had been males. All three retired within two years, leaving only females among the veteran English teachers.

Hester Hypoint had taught English for nineteen years, the last seventeen of those years with the ninth grade. Cheerful and determined, she was a well-respected member of the school community. Because her teaching schedule consisted entirely of ninth grade classes for so many years, she could fairly claim to have taught nearly every graduating senior for two decades. She was also in a position to compare the students over the course of these decades and give her opinions as to their overall preparation for the high school. Her long tenure with the same grade level was typical of the veterans. Any "experimenting" with other grade levels had taken place early in their careers, and

each had spent the major part of their careers working with students at the same grade level or levels.

Jeanette Jones had taught junior high English, grades seven and eight, for over twenty-five years. She had always taught at the junior high level. Like her veteran peers, she had long experience with the same level and felt confident to compare incoming classes with those she had taught in previous years. She had become head of the English department a few years prior to the study. One of her chief concerns in dealing with the seventh graders she instructed was that they should become "good school citizens." Among other classroom activities, she ran the Thanksgiving food drive for the poor and awarded extra credit points for bringing in canned goods during the holiday season.

Karen Kistner was a thirty-year veteran who had always taught remedial English and reading to grades seven through ten. Although her work had always required working with students who were typically labeled as "under-achievers" or "low performers," she expressed concern that the overall performance level of the general student population had declined steadily—if in nearly imperceptible ways—over the thirty years of her career. She questioned the usefulness of testing regimes for students of the kind she usually worked with. Because they were often far below the grade level they found themselves in, holding them to testing standards that matched with the other students who were actually at grade level seemed to lack sense and "might even be cruel."

Luanne Lender had taught in both the junior and senior high over the course of her twenty-seven years of experience. Her acquaintance with the district went back several years beyond the twenty-seven, however, because she had left school teaching to raise her children, returning to the school after a hiatus of eight years. She had retired just

before the study was begun, and had spent the last eighteen years of her tenure in the district teaching eleventh grade English. She had taught both the "regular," or average, track and the honors track for many years. In the final years of her tenure at the school, the honors track was eliminated on the grounds that it was largely a social rather than academic honor**;** she felt that this was a sign that excellence in English was in serious decline. "I was really concerned when the honors classes were eliminated," she said. "There's really no place for the gifted student in the system anymore."

Mary Minton had also retired in the year before the study began, and had taught in the same district for thirty-three years. While she had taught grades seven, eight, nine, and ten, she had spent most of her years teaching seventh grade, and "liked seventh and tenth grade best." She had served as department chair for the last seven years of her career, being succeeded by Jeanette Jones. She supported the elimination of the honors track, arguing that it did not promote academic achievement, but had become "a special club for a small group of students" who, having been placed in eighth grade honors English, began to "coast," and were no longer advanced compared to the general population by the time they reached tenth grade, "but you were never going to get them out of the honors class." She did not consider the elimination of the honors track in itself a sign of general deterioration, although she agreed with Mrs. Lender that the overall performance of students in general—in all subjects "and probably in most schools"—was in decline. All of the teachers who participated in the study are listed in the following table.

Teacher Name	Grade level(s)/courses	Experience/Education
Alice Andrews (novice)	10, 12 honors, journalism	1.5 years / B.A. (English)
Betty Browne (novice)	7 and 8	2.5years / B.A. (English)
Catherine Carney (novice)	9, 12, 11/12 film elective	3 years / M.A. (Eng.lit)
David Dutcher (novice)	9, 11, 12th public speaking	5 years / M.A. (Eng.lit)
Ellen Enders (novice)	10, 11	5.5 years / M.A. (Eng)
Fiona Fisher (established)	7, 8, 10, 11	8 years / M.A. (Educ)
Frank Fender (established)	10, 11, 12	8 years / M.A. (Eng.lit)
Gail Goodwin (established)	11, 12, 7, 8	12 years / (M.A. Eng.lit)
Hester Hypoint (veteran)	9	19 years / (M.A. Eng. lit)
Jeanette Jones (veteran)	7, 8	25 years / (M.A. Reading)
Karen Kistner (veteran)	7, 8, 9, 10	30 years /(M.A. Reading)
Luann Lender (veteran)	11	31 years / (M.A. Eng.lit)
Mary Minton (veteran)	7, 8, 9, 10	33 years / (B.A. English)

Data Collection

Interviews with each of the teachers were conducted over the course of an entire

school year (fall 2001 - summer 2002). The interviews were semi-structured (Sherman &

Webb, 1988) and designed to encourage the participants to explain their approaches to grading, both in terms of the practical issues of weighting and balance among assignments and, in later interviews, in terms of fairness and perceptions of fairness. The interviews were tape recorded with the consent of the teachers, and significant portions of the recorded interviews were transcribed in order to allow for analysis. As a form of member-checking (Stake, 1995), the teachers were invited to review the transcriptions for accuracy.

In the subject school district, reporting of grades is driven according to the calendar, with four marking periods of ten calendar weeks (i.e., not ten weeks of classes). The interviews were conducted so that they coincided with the report card periods.

In addition to report cards, the school issues progress reports halfway through each marking period. These consist of a computer-generated report which allows the teachers to enter a grade and up to three comments from a list of pre-scripted remarks. This format replaced a long-standing tradition of sending "five week letters" between marking periods. Letters on blue paper commended students who were progressing satisfactorily; letters on white paper (produced in triplicate, with one copy for the parents or guardians, one for the classroom teacher, and one for the guidance department) warned of deficiencies in performance or behavior. Interviews were also scheduled at the five-week periods, since the progress reports involved evaluation decisions, and certainly raised moral considerations.

Every teacher could not be interviewed at each of the five-week periods, but in every period, most of the teachers were interviewed, and those who were not were asked the same questions later, when their schedules again permitted them to meet with me. The

purpose in having several interviews at several times over the course of the year was to allow for comparison among the teachers, across the interviews, and over time. This allowed for triangulation by source and time, and provided a means of judging consistency in, and/or evolution of, grading practices over the course of the year.

Initially, moral issues involved in grading were examined through indirect questions in order to keep the participants at ease and elicit frank responses. Because of the very real danger that a threat to the validity of the data could result from a response set (Gay & Airasian, 2004), teachers were not asked direct questions about "fairness" and character development issues until well into the school year, by which time they seemed comfortable with me and the tape recorder. Several questions (see Appendix 1 for representative questions) were designed to reiterate earlier questions following discussion in order to see if individual responses varied with changes in context or over time. The purpose was not to "catch" the teachers in a contradiction, but to allow for triangulation of the data, and to see if changes are made to grading policies and how such changes are justified in the minds of the teachers. Decisions about grading are bound up in context and circumstance, and the year-long collection of data allowed for an examination of how the teachers adjusted their practice as their students—and the school year—progressed.

Interviews were scheduled during the school day and lasted between forty minutes and an hour each. The interview schedule was adjusted as required; most of the interviews occurred within a week of the end of a marking period or warning period. The teachers were encouraged to clarify and expand upon answers to questions, to give examples of both typical and unusual grading decisions, and to explain their thinking in making decisions about assigning grades.

The high school principal was also interviewed in order to obtain information about grading practices at the high school and within the English department. He provided little concrete information, stating that there was no set policy on how grades were to be determined, and that he believed it "might be a good idea to have a discussion" with the English teachers about grading and grading practices. In general, he believed each teacher had "figured out how to grade," and that the department chair and other teachers helped the newer teachers with suggestions or advice as required.

Documents Analyzed

While interviews comprised the bulk of the data, documents involved in grading were examined. These included the teachers' handbook, copies of the official "blue and white letters," copies of the five week progress report, copies of the report card grading forms, and the accompanying "comment sheet" (the report card contained a field for three brief "standardized" comments such as "needs to pay more attention," "fails to do homework," "needs extra help," "a pleasure to have in class," and so on. These were to be chosen from a list of nearly a thousand comments, each with its own numeric code. The code would be entered, or "bubbled in," in the field provided). Grade books kept by the teachers were examined when they were useful for providing additional information. Field notes were kept along with the interviews (Bogdan & Biklen), and these notes were used to identify the context and time period as interviews were compared with one another.

Human Subjects Committee Approval

Human Subjects Committee approval was obtained for the study, and all interviews took place during the 2001-2002 school year, with some follow-up data collection

continued as needed through December 2002. All rules imposed by the Human Subjects Committee were observed throughout the study. The Study Information Sheet from the application is included in Appendix 2; each of the participant teachers received a copy of this form prior to being included in the study. Interviews were taped with the consent of the participants, and the tapes coded to maintain confidentiality in keeping with the protocol established for Human Subjects Committee approval.

Data Analysis

The data collected were transcribed from the interview tapes, sorted and categorized, and compared with the field notes, following a process of continuous comparison (Glaser & Strauss, 1967; Strauss & Corbin, 1994). During the initial sorting, major themes emerged and the main similarities and differences in grading practices among the participants became clear. These were summarized, and the data summaries were shared with the participants for validation (Stake, 1995). This member checking also served to engage the participants in further discussions, so that issues beginning to emerge from the analysis could be further explored. Next, the data supporting the analysis were grouped to facilitate comparisons among the teachers, and tentative categorizations were formulated (Ryan & Bernard, 2000; Taylor and Bogdan, 1984). Data collected in "feedback interviews" were then used to further delineate the issues involved.

Six preliminary areas emerged as topics for analysis and examination in terms of moral implications. The philosophical framework derived from Jackson et al. and described earlier served as the means for examining these six areas and analyzing the moral issues which arose when comparing the teachers' practices bound up in each of the areas.

While the process of grading is a complicated and dynamic one, the initial examination of each of the six areas separately allowed for "teasing apart" the many strands of complexity. Treating each separately also allowed for a more coherent review of the pertinent literature touching upon each area. An attempt at carrying out a "comprehensive" review of literature on the topic of grading would have only obscured the implicit moral issues; by separating these six areas and examining each in isolation, the moral aspects could be more clearly articulated.

The six areas which emerged after the initial analysis were these:

1. Grading Systems - the ways in which value is assigned to school work and relative weights are given to different assignments, and the ways in which these weighed assignments are translated into report card grades;

2. Subjective judgment - how work, especially writing, that requires subjective grading is evaluated;

3. Rubrics – The bridge between the first two areas; a device for applying the grading system to both processes and products and turning abstract features like "understanding" and "expression" into concrete points or grades;

4. Dealing with Effort – how effort is perceived and whether teachers use grades to reward the diligent and punish the lazy, or to encourage the weary;

5. Forming Character- a process submerged in all of the areas, and often not fully perceived or acknowledged;

6. "Fitting to the System"- how teachers adjust their grading practices to fit the context of the school system and the expectations of the wider school community.

Further sorting and grouping of features within these groups led to a reshaping of the areas into chapters; subjective judgment, rubrics, and "fitting to the system" were subsumed by the larger category of grading systems, which constitutes a single chapter.

Because the teachers spoke about effort in ways that showed that the concept was almost inextricably tied up in judgments about attitudes and the teachers' expectations for each student, the theme of effort was expanded to become "effort, attitude, and expectations," which constitutes another chapter.

Character development, although related in many ways to the teachers' feelings about effort, attitude, and expectations, was in fact a pervasive implicit theme, and so it also forms a separate chapter.

Each chapter begins with a literature review centered on the area being considered and providing context. The literature review is followed by a comparison of the department members' ways of dealing with each area. A process of inductive analysis is used to draw conclusions about the information obtained. Support is generally presented in the form of quotations taken directly from the data. Negative case evidence which seems to contradict the main body of evidence is considered carefully and presented along with the main analysis. Finally, the moral implications of the actual practices are examined using the philosophical framework. The final chapter presents findings about the grading process as a whole and its moral implications.

CHAPTER THREE: GRADING SYSTEMS: FINDING A BALANCE

Every teacher must determine a system of grading which makes it possible to take into account the various kinds of work the students submit for the teacher's review, and to communicate the teacher's understanding of the relative worth of any given assignment in relation to other assignments, to previous assignments of a similar kind, and to the final outcome of the course. How does one compare a pop quiz on a reading assignment to be done for homework with a unit test given at the end of a novel? How does a student know which assignments are "more important" and which less? The way a teacher decides to weigh—or not to weigh—the many different kinds of assignments given over the course of the school year affects the students' understanding of their grades, and the system a teacher uses for assigning values to particular assignments has important moral implications.

This chapter begins with an exposition of information regarding methods for setting up a grading system and underlying issues as found in the current literature, and then presents the grading systems used by the teachers of Mellmax High School, with some exposition of the ways their chosen grading systems are used in actual practice. After the data is presented, an analysis using the theoretical framework established in chapter one allows for a discussion of the moral implications of the features which characterize each system and the ways in which the teachers actually implement their chosen systems.

CONTEXT AND LITERATURE

Definition of a Grading System

As the term will be used here, a "grading system" is a method for converting different kinds of assignments into one kind of grade so that a single composite grade (usually a

mathematical average of all the scores given) which somehow captures the general classroom performance of an individual student over a period of time can be entered on a report card. While this seems like a simple enough concept, it is in fact an extremely involved undertaking. The literature on grading reveals a history of swings from one type of system to another, and back again (Brookhart, 2004; Kirshenbaum, Napier, & Simon, 1971) as teachers struggled with the many kinds of things that can be graded, the types of information schools, parents, students, and the teachers themselves believe are important or worthwhile, and a way to condense the constant formative and summative assessment of the classroom into a fixed quantity at regular intervals in report cards.

Grading Reflects Purpose and Philosophy

The purposes of grading underlie the entire grading system, and researchers, school administrators, parents, teachers, and students all have their own ideas about what purposes grades are meant to serve. Grades are the result of a teacher's scoring of a test or assignment. Karmel (1970), citing Webster's New Collegiate dictionary, asserts that "the use of *test* in education means 'any series of questions or exercises or other means of measuring the skill, knowledge, intelligence, capacities or aptitudes of an individual or group'" (p. 4). Using this broad definition, virtually every assignment a teacher gives is a test, and according to Karmel,

> ". . . schools use tests as educational tools to promote individualized
> instruction. . . [this] implies that the school's basic duty to the child is to
> know him as an individual. Inherent in this is a recognition of the dignity
> and worth of the individual and his unique qualities. The basic premise for
> giving tests is the assumption that individuals differ and that education

52

must be geared to these differences so that each person may develop his or her own unique potential" (p. 4).

Karmel's notions of the purpose of grading reflect one of the major philosophical strands present in American public schooling and resonate with the whole child movement of the 1940's. They are by no means out of date; however, they may not reflect the current dominant viewpoint regarding the purposes of grading, which is perhaps better summed up in a more clinical analysis by Brookhart (2004):

> "Current dilemmas about grading are seen most clearly as a confusion of purposes. . .grades have been used to serve three general purposes simultaneously: ranking (for sorting students into higher education, for example); reporting results (accounting to parents the degree to which students learned the lessons prescribed for them); and contributing to learning (providing feedback and motivating students)" (p. 23).

Stiggins and Conklin (1992) discuss assessment purposes in light of their own study of classroom teachers. They list the following "purposes for assessments":

> "diagnosing individual and group needs, sizing up students, grouping and placement, assigning grades, feedback to students, parents, and school managers, control and motivation, communication of expectations, instructional decision making, and preparing students for later assessments" (p. 57).

However, they go on to report that "two of the three classrooms we observed required that assessments be reported out in the form of [number or letter] grades" (p. 57). Since

the report card format used by the teachers of Mellmax High School allows for only a numerical grade and a handful of pre-scripted comments, it is clear that the major part of assessment at Mellmax High is also aimed at arriving at grades, and that this results in a number that reflects a composite of many different assessments.

The differences among authors regarding the purposes of assessment and grading is a reflection of the way the pendulum swings regarding grading as differing philosophies about the purposes of school jostle for dominance. Kliebard (1995) traces the mood swings of the American public schools over nearly three-quarters of a century. He shows that while debate over the purposes of schooling and ways to measure the accomplishment of those purposes constantly shifts the emphasis from one area to another, it is never settled. The temporary stifling of one purpose (say, individual development) in favor of another (say, equal performance against a given standard) goes on only until it is apparent that the weakening of one purpose is detrimental to the whole. At that point, the emphasis shifts and the neglected component is reasserted, usually in a way that stifles some other component. And so the cycle continues.

Shifts in philosophy regarding the purposes of school are only part of what influences the balance of grading and judgment in school. Debate over what can be measured, how it can be measured, and whether it should be measured at all adds another dimension to the grading question. Brookhart goes on to explain that the "confusion of purposes" she detects in schooling results from the need for different methods of referencing in order to construct measurement scales for each purpose (i.e., norm-referencing for ranking, criterion-referencing for achievement, and self-referencing for judging personal development). Brookhart then asserts that, "According to current measurement theory,

this is a recipe for disaster" (p. 23).   Her acknowledgment of issues raised by measurement theory signals another source of debate.

Two Camps: Measurement Experts and Classroom Teachers

An interesting historical fact is that the rise of the American high school occurred at nearly the same time as the growth and development of widespread interest in testing and standardization of tests to compare and classify individuals in relation to others or against some criterion of interest. The number of high schools in the United States multiplied nearly exponentially in each decade from 1860 to 1900 (Harris, 1901), and notions of "social efficiency" and the "scientific curriculum" came together in the high schools (Kliebard, 1995). Despite the fact that measurement on a grand scale could reveal little about a teacher's satisfaction with the day-to-day performance of his or her students, the temptation to transfer techniques for mass testing to the classroom was great, and continues to influence the way testing and evaluation, even on the small scale of the local classroom, is understood.

Interest in ways to sort, classify, and compare individuals, especially in terms of career-worthiness (Kliebard, pp. 77-130), swelled as schools became a fixture in the American educational landscape. The pioneers of intelligence testing, people like E. L. Thorndike, Francis Galton, Wilhelm Wundt, and Alfred Binet, were all working in different countries but guided by the same notion that mental testing would provide a means of accurately determining an individual's actual intelligence. All these "fathers of standardized testing" were interested in the intersection of schooling and mental development. Classic studies of intelligence and the effects (or lack thereof) of schooling on intelligence were conducted in the early days of the twentieth century (Thorndike and

Woodworth, 1901; Thorndike, 1906; Thorndike, 1924), as the high schools took root in the American scene.

One of the important outcomes of this confluence of scientific interest in measurement, especially of intelligence, and of the debate of the purpose and direction of the public schools is the confusion that results when discussing grades and grading practices. The discussion between the "measurement community"—as Cross and Frary (1999) call the psychometricians, measurement experts, and standardized test makers— and the everyday classroom teacher about grades and grading had already begun in the early days of the high school, and continues today. Often this results in each side treating the other with suspicion; the measurement experts see the classroom teachers as careless graders at best, and capricious ones at worst, whereas the classroom teachers dismiss the experts as meddlers who don't understand the realities of the classroom.

The result of this ongoing discussion is a public discourse concerning grades and grading systems that suffers from a failure to distinguish the fundamental differences in the objectives of the two camps, and causes confusion among all interested parties—the public, classroom teachers, school administrators, and the measurement experts. Even as teachers worry about validity, reliability, and other concerns close to the hearts of psychometricians, they are concerned with "the need to manage classrooms and motivate students" (Brookhart, 1994, p. 299), develop character, and maintain a pleasant working atmosphere. Finding a grading system that enables all of this is a challenge; perfecting one that meets the teacher's needs and conforms with his or her philosophical inclinations about instruction is a career-long process.

In the midst of all of this balancing and re-balancing, the classroom teachers move along in their daily assignments, not entirely oblivious to the concerns of the measurement experts, but not unduly influenced by those concerns, either. They continue to work out a balance of assignments and grading that they believe achieves the purposes of their course and school.

The Challenge of Finding a Grading System that Works

Finding a grading system that matches the purposes a teacher has in mind and is at the same time simple enough to be easily managed has been one of the great challenges of schools for centuries. Brookhart (2004) summarizes typical grading schema from 1640 until the present (pp.15-27), detailing "major developments" in grading and revealing the fundamental difficulty in arriving at the "perfect" system.

Although Brookhart's history is little more than a sketch, it establishes that from the beginning, grading practices were based on the assumption that "it was the duty of the faculty to evaluate students, and that the merits of students could be estimated rather precisely" (p. 15). Grading in the universities depended upon examinations, "which were sometimes conducted on the day of graduation itself" (p. 16), and over time led to distinctions among students. The first recorded scale for grading in the United States was a set of categories used at Yale in 1785: "Optimi, second optimi, inferiores, and pejores" (i.e., best, second best, lesser, and worse). The four-point system was improved upon (again, by Yale) in 1815, where numbers from 1-4 were introduced and decimals allowed for finer distinctions among students (p. 17). This scale is still used in most colleges and universities, where college professors struggle in ways similar to high school teachers to

capture their true opinion of a student's standing in ways that are seen as equally dubious in the eyes of the measurement community.

By the time report cards appeared in the common schools of the 1800's, teachers were already complaining about the burdensome nature of preparing the reports, and parents about the difficulty in interpreting them. Brookhart quotes a contributor to *The Common School Journal* in 1840, identified only as "S.G.B." who describes the report card of the day: "In some schools, the practice has been adopted of using printed forms, containing blanks, in which, by some system of figures and letters, the advancement and behavior of the pupil are to be expressed by the teacher" (S.G.B., 1840, in Brookhart, 2004, p. 18).

Essentially, "S.G.B." describes the report card of today. In the intervening years, however, the "figures and letters" were sometimes in the form of *percentage grading*, from 0-100, *letter grades*, A through F (with or without plusses and minuses), *pass/fail*, or *narrative grading*, where teachers wrote comments on the cards in the spaces provided. As various ways of thinking about schooling gained prominence, the system that seemed to capture the current mood best was used, only to be replaced or modified as the thinking changed. Each of the systems seemed at first to be the ideal answer to the grading question, yet each was soon found to be inadequate, or at least, less than perfect. This was further complicated by the fact that grading was (and still is) largely left up to the individual teacher—who, after all, would be the one person in the position to evaluate the students in his or her own classroom—and each teacher created a system to suit his or her own purposes, observing, of course, whatever external constraints were imposed by the school district. Odell (1925) found that in Illinois alone, teachers in the public schools

used over a hundred different grading systems, and concluded that apparently similar grades might be the result of very different grading approaches. Despite efforts to standardize grading, teachers still customize their grading schemes, for the most part, as they see fit.

Doubts about Teacher Grading Reliability

The use of percentages in the common school was seen as an advance over the four-point system, especially since it gave a greater range of scores, and, supposedly, an ability to make finer distinctions in judgment. Use of percentage grading dominated the common school by the end of the nineteenth century, but shortly after the turn of the twentieth, studies such as those of Starch and Elliot (1912, 1913a, 1913b) demonstrated that the reliability of such grading was poor.

Starch and Elliot's classic experiments asked many different teachers to grade the same paper. Using an English essay first, they found a range of thirty-nine points on a scale of 0-100, with a passing score of seventy-five. When critics complained that it was the subjective nature of the English essay that accounted for the wide range of scores, Starch and Elliot conducted new studies in math and history, and found that the range in grades on a geometry paper was nearly fifty points! The history paper suffered equally at the hands of its graders. These studies are often interpreted as "a landmark in casting doubt upon the reliability of testing and grading procedures" (Kirshenbaum, Napier, and Simon, 1971, p. 259), but they are perhaps more important in that they demonstrate that grading is not so much a function of the test or its subject matter, but of the grader and the method of testing, i.e., *who* gives the grades and *how* he or she gives them needs to be known in order to actually understand what a grade means. In any event, this first clash

between the measurement community and the classroom sent interested parties scurrying to find a way to limit the difficulties with percentage scores and reliability.

From Percentage Scores to Letter Grades

As researchers (Pressey, 1925; Odell, 1925; Bells, 1930; Sims, 1932; Adams, 1932; Dexter, 1935) continued pursuing the reliability of grading issue, schools turned to the letter grade, which allowed a teacher to capture the worth of a paper within a range, thus seeming to reduce the reliability problem. If an A meant a score between ninety-two and one hundred, the need to distinguish between a ninety-four and a ninety-five vanished. However, the difference between a B and an A could still be a problem, and teachers still felt a need to distinguish between a really solid B and a shakier one, so "plus" and "minus" were eventually added to increase, once again, the ability to make finer distinctions within grade ranges. Whether changing from a finer scale to a coarser one actually does anything to improve reliability, rather than merely masking unreliability, is doubtful, yet letter grading dominated the schools in the 1920's and 1930's (Brookhart, 2004, p. 16).

Studies of grading reliability in subsequent years (Tieg, 1952; Kirby, 1962) continued to demonstrate that grading across teachers is unreliable, as did similar studies in universities (Bass, 1951; Thompson, 1955; Aiken, 1963; Temple University, 1968; University of California at Berkeley, 1965). The conflict between a perceived need to distinguish closely among students (requiring percentage grades) and a desire for improved reliability meant that grading schemes alternated between percentage grades and letter grades.

Pass/Fail and Mastery Grading

The University of Michigan, in an attempt to do away with the need for drawing any distinctions beyond competence vs. incompetence, experimented with the pass/fail grade in 1851 (Brookhart, 2004, p. 24). This reduced the necessary decision to one of mastery alone—had the student learned enough to pass?—and may have contributed to the mastery grading schemes of the mid-1920's (Brookhart, 2004, p. 24). Bloom, Hastings, and Madaus (1971), discuss the strategy of mastery learning, which comes down to a pass or fail grade.

The chief difficulty of mastery learning is the requirement of different pacing for different students; in high school, this is generally seen as unworkable due to time constraints and the difficulties that can arise if (a) a few students far outstrip their classmates in mastering new materials, but cannot be advanced into another classroom more in keeping with their skills, or (b) a few students take so long to master something that the entire class is being held back as the teacher directs an inordinate amount of attention toward these weaker members.

Some students, especially special education students mainstreamed into regular classrooms, may be graded pass/fail, in part to prevent a feeling of stigma associated with low grades. Whether a "pass" in such cases indicates "mastery" is doubtful, and raises the concern that the broad range encompassed in the "pass" grade includes not only true mastery, but also minimal—and perhaps only temporary—competence.

Often teachers grade homework or minor assignments on a pass/fail basis and then assign some overall value to the "homework grade" which is based on the amount of homework completed. This is not so much a judgment of the quality of the work as it is a

judgment concerning attitude and effort through inference—i.e., regular completion of homework indicates a positive attitude and provides evidence of effort.

Narrative grading

Many kinds of assignments require feedback in the form of commentary if students are to profit from assessment, and teachers struggle to put into words their evaluation of assignments that require subjective consideration. Sometimes, teachers feel that they must add commentary to a grade in order to justify it—perhaps with a remark about a decline (or improvement) in performance compared to previous work. At other times, depending upon the student, a teacher may feel that it is necessary to "soften the blow" when a grade is likely to seem harsh, or if it is likely to draw criticism from the student or the students' parents. Narrative grading is often used, then, in conjunction with percentage or letter grading (Brookhart, 2004).

School districts sometimes use a computer-generated report card that provides a "menu of comments from which the teachers can select; sometimes handwritten report cards have sections for teachers' comments; and sometimes teachers, either because of district policy or on their own, send home separate narrative progress reports" (Brookhart, 2004, p. 27). The Mellmax school district uses a computer generated report card that allows for a letter grade and up to three comments selected from a list; the five week progress report is similar, except that the grade is given as a range in ten point intervals; the list of comments is the same. While this is the extent of narrative grading on the report cards, all of the teachers in this study declared that commentary on assignments involving subjective judgment was their everyday practice.

Evaluation, not mere measurement

Typically, texts written to help teachers learn about grading make a distinction between "measurement" and "assessment/evaluation." Horrocks and Schoonover (1968) make the distinction thus: "Ordinarily 'measurement' is the preferred term when tests are used, 'evaluation' or 'assessment' when more subjective judgments are used" (p. 3). This distinction makes *test* a much more specific term than it is in Karmel's writing discussed earlier. Yet Horrocks and Schoonover follow up on this distinction by observing that "the dimensions of human behavior and capacities with which measurement is most directly concerned include intelligence, special abilities and aptitudes, personality and temperament, attitude and opinion, achievement, and social behavior" (p. 3).

The inclusion of such items as temperament and attitude implies that many of the things typically thought of as subjective on the part of the classroom teacher—and therefore "evaluated" rather than "measured"—are simply not yet being measured using the correct method. However, Horrocks and Schoonover concede that ". . . sound measurement can proceed only when substantial knowledge exists regarding the nature of the thing that is being measured. At this time, some of the variables of behavior have defied analysis and in the present state of knowledge are unmeasureable' (p. 4). This highlights the dilemma of the schoolteacher in evaluating students' overall performance. If the students must be assessed, but some of the elements of such assessment are, in scientific terms, "unmeasureable," how does a teacher balance those elements into an equation which demands, in the end, a fixed, quantified grade?

Karmel explains that teachers use tests (and thus, grading) because

"in order to gauge . . . progress, the teacher must institute evaluation techniques. These techniques include essay and objective tests and informal procedures, such as day-to-day classroom observations and teacher judgment based on professional experience and intuition. All of these procedures aid the teacher in evaluating pupil progress" (p. 5).

Objective and Subjective Grading

Addressing the issue of "objective" and "subjective" grading, Karmel asserts that "*Objective,* as used in testing, means that the scoring is not influenced by the opinion, knowledge, or skill of the person scoring the test; or whether the person taking the test and the person scoring the test 'communicate.' In. . . essay questions the correct answer is subject to interpretation by the teacher. The scoring, is, therefore, subjective' (p. 5).

It is easily accepted that the English teacher is an appropriate judge of the quality of a written work, both in terms of the mechanics of the piece (spelling, punctuation, grammatical structure, and so on) which are objective, and in terms of style, usage, effectiveness of structure, and so on, which are subjective, because he or she knows about such things and can be trusted in terms of his or her judgment. But there are many other classroom activities which the teacher judges that are not part of the knowledge base of English, writing, or literature—things like attitude, diligence of application, cooperation, and so forth. Judging these things is highly subjective, and no doubt it is for this reason the measurement experts shy away from them.

But even the objective parts of grading may not really be as objective as they seem to be. As the movement to create standardized school tests like the SAT's gained momentum, Rothny (1955) observed:

"the so-called 'objective test' is really a subjectively constructed test that is objectively scored. The actual writing of the test items is a subjective process. The author of an objectively scored test must decide on the materials he will sample, must make judgment about whether or not an item is worthy of inclusion, and must select among scoring schemes . . . no objective scoring system can ever make up for faulty subjective decisions made during the construction of the test" (p. 13).

In the end, the many decisions a teacher must make about what to test, how to test, and even whether to test are subjective, and it is perhaps this fundamental reality about grading that makes it a morally-laden undertaking.

Achievement and Non-achievement Factors

Schooling involves a number of factors which play into the mission of the public school. High school students are growing adolescents approaching adulthood, and schools are expected to contribute not only to the academic advancement of these young men and women, but also to their development as thinking, productive, responsible members of the wider society which they will join after graduation. As a result, "assessment" becomes a tangle of decisions and judgments, and the chief concrete expression of this assessment is the report card. In a broad sense, the "grade" a child gets is not merely the numerical grade, but a combination of the numerical grade and the commentary, as limited as that is at Mellmax High. The grading process involves not only the calculation of the numerical average of all the assignments the teachers have scored, but a careful consideration and selection of the comments to be included on the report card.

However, decisions on promotion and advancement are almost always based on the numerical grade alone, so that even comments that "raise red flags" about the student's social failings or anti-social tendencies (or are intended to suggest greater potential or value in the student than the numerical grade seems to indicate) have little to do with the decision to pass or fail a student. This may encourage teachers, even if not in a fully conscious way, to "transfer" some part of the evaluation from the commentary into the numerical grade, as a kind of attempt to prevent the passing—or failing—of students who seem to deserve different treatment if judged according to the fields addressed in the commentary rather than according to the numerical grade alone.

In part as a result of the difference in perspectives between the measurement community and the classroom teacher discussed earlier, it is common to find admonitions in texts on grading about which factors should be "graded" and which "assessed" but not graded. Several authors (Ebel and Frisbie, 1986; Anderson and Bourke, 2000; Stiggins, 2001; Brookhart 1994, 2004; Hopkins et al., 1990; Gronlund, 1985) assert, in varying degrees of vehemence, that *only* achievement should be weighed in determining grades, and "non-achievement factors" such as interest, attitude, effort, attendance, deportment, motivation, personality, and social interaction with others should not be allowed to enter into grading calculations.

The distinction may be finer than most teachers will make, and while most of the teachers interviewed for this study declared in early interviews that certain "non-achievement" factors were not used in grading, some admitted some such factors were, and even those who were more scrupulous about excluding such factors at the beginning

of the year allowed that, as the year went on, some of the non-achievement factors did indeed creep into some grades.

Other authors (Airasian, 1996; Airasian and Jones, 1993, Wiggins, 1994, 1997) are more tempered in their considerations of the demands school places on teachers with regard to assessment, and have suggested that the measurement community broaden its parameters regarding appropriate classroom assessment and grading practices to include at least some of the day-to-day assessments teachers regularly make for managing their classrooms and making instructional decisions, echoing Karmel's endorsement nearly twenty-five years before of informal assessments based on "professional experience and intuition."

One might question the construction of the term, "non-achievement factor," since it is apparently meant to indicate an influence on something (the teacher's perceptions, perhaps?) that is not part of a student's achievement. But it also suggests a factor that *has no influence* on achievement. Ignoring, for the moment, the list of factors defined as such above, one could ask, "What is a non-achievement factor?" What you had for lunch, perhaps, would not be a factor in achievement. But is effort not a factor in achievement? Granted, the measurement of effort is imprecise and teachers can be fooled by their students as to how much effort they actually put into a given assignment, but is effort properly a "non-achievement factor" to be excluded from the grading scheme on the grounds that it is hard to measure?

Participation is another "non-achievement factor," as is attendance. But one can reasonably infer that participation both aids and indicates learning, and lack of participation—or the kind of participation that actually interferes with instruction—limits

learning and gives little indication of learning, making it harder for a teacher to tell if a student is advancing as hoped for at the given moment. And without attendance, participation simply doesn't occur. Many school districts cite research that correlates attendance with test scores, suggesting that there is a positive relationship between attendance and achievement, so incorporating attendance into a grading scheme does not seem altogether unreasonable.

Hodgepodge Grading or Multivariate Analysis?

Because the admonitions of the measurement community regarding best practices in grading "are generally dismissed by teachers as 'unrealistic,' 'impractical,' or more bluntly, 'not relevant to classroom needs'" (Airasian and Jones, 1993, p. 241), teachers find themselves with a broad array of grades for a wide variety of things which are sometimes adjusted from student to student. This results in what Brookhart (1991) calls a "hodgepodge grade of attitude, effort, and achievement" (p. 36). Parsons (1959) had earlier observed the widespread use of a combination of factors in arriving at grades, but had more kindly classified the components into two categories, observing that students were "defined in terms of a fusion of the cognitive and moral components, in which varying weights are given to one or the other" (p. 304).

When discussing grades and report cards, Karmel acknowledged the importance of multiple sources of data—tests and quizzes, projects and papers, reports and essays, and even classroom participation. Regarding some of the less concrete data sources, like participation or attitude, he suggests that "probably it is best not to include them in overall grade evaluation. They should be noted in the child's cumulative record, and if there is space on the report card for comments, a brief written description and analysis of

these traits should be given" (p. 423). He further suggests that report cards should include spaces for commentary so that achievement and other kinds of measures can be separated. In the end, however, he concedes that "one must face reality and admit to students and parents that grades are only an attempt to evaluate performance" (p. 425). "Performance," however, is a broader concept in a schoolteacher's mind than "achievement," which may be what Karmel actually means; if a grade is an attempt to evaluate performance, then behavior in the classroom is an important element in a child's performance, even if it is "unrelated" to achievement.

Surveying 307 teachers and 8,664 students in a single-school system, Cross and Frary (1999) set out to confirm previous research that supported Brookhart's claim that teachers' grading practices were a hodgepodge of disparate elements.

> "The results largely validate the findings of earlier studies. Substantial majorities of the teachers reported 'hodgepodge' grading practices. More important, the students largely confirmed and supported the hodgepodge grading practices reported by their teachers. These results are contrasted with grading practices widely recommended in measurement texts and are followed by a discussion of how measurement specialists may be missing the mark in their efforts to communicate their views to teachers, school administrators, and the general public" (p. 53).

While Cross and Frary are suggesting that the problem is that the measurement community hasn't made its case well enough, one could counter that the trouble lies not in ineffective communication, but in an unwillingness to accept the wide variety of factors that those directly involved in the actual day-to-day activities of school believe

are important enough to be weighed into the grades that are assigned. To use the jargon of the measurement experts, perhaps the problem is that the means for accurately measuring the less tangible non-achievement factors have yet to be developed so that a scientific "multivariate analysis" can be carried out. It seems unlikely that the teachers will abandon their multiple sources of information, whether this looks like an indecipherable hodgepodge to those outside their system or not.

The role of the report card

Wiggins (1994) suggests that one resolution of the problems with the wide variety of sources for grades might come from better-designed report cards. He proposes six "new" approaches for designing report cards:

1. a clear distinction between standard-referenced and norm-referenced achievement in reports;

2. a system that sums up teacher judgments about progress toward exit standards and about growth with regard to teacher expectations;

3. a longitudinal system to compare students over several years;

4. many more sub-grades of performance;

5. distinctions between the quality of work and the degree of difficulty of the work;

6. evaluation of "intellectual character"—"habits of mind and work based on performance and products" (p. 28-29).

Whether this would be a practical solution or result in early teacher burnout due to the burden of reporting requirements, most schools strive to strike a balance between giving enough information and ease of reporting. The teachers at Mellmax High School take the

nature of their report card into account when determining their grades, and feel compelled to work to transform a hodgepodge of material into a detailed analysis which is then expressed in a single grade and a handful of pre-scripted comments.

Classroom factors that don't look like assignments—but are

It is natural to assume that a grading system would reflect the intended purposes behind giving assignments—to demonstrate understanding, to prove that something has been memorized, to practice a skill, and so on. Some assignments may be aimed at determining underlying attitudes or developing thinking abilities. And some assignments—or tests, if you will—may not at first be apparent as such. For example, if a teacher explains the "classroom rules," following the rules becomes an implicit ongoing assignment. The day-to-day adherence to the rules, then, can be seen as a successful completion of this assignment. Failure to follow the rules is failure in this assignment. While the teacher may not be marking down each infraction of the rules by each student, he or she is keeping a kind of mental gradebook, and at least the students at the two extremes (i.e., the best behaved and the worst behaved) are likely to find that the teacher has used this information in weighing out grades.

It is perhaps these purposes that are less visible that confound the grading debate—in part because, as we will see, teachers do indeed test their students against a standard of conduct and effort that is not part of the daily exchange of "gradeable" work, but have no pencil-and-paper method of measuring such things.

Arriving at a fair grading system depends in part on the perceptiveness of the teacher and the teacher's ability to judge fine differences in performance. It also depends on the correspondence between what is assigned, whether explicitly or tacitly, and what is

actually desired as an outcome, whether social or academic. When what is assigned cannot be used to actually "prove" that some objective has been achieved, then whatever is actually being judged becomes uncertain. This raises the classic questions of validity which surround testing of all kinds—is the teacher actually testing what he or she plans to test? How does one deal with a grade that is inconsistent with most or all of the previously obtained grades? How can these grades predict future performance?

All of these considerations affect the kind of grading system a teacher decides to put into place in order to change the tasks assigned over the course of a school year into the single grade that sums up a marking period, a semester, or the entire year. Conversely, however, once a grading system has been settled upon, it can become the engine that drives assignments, so that it can *determine* classroom practice rather than being shaped by it.

GRADING SYSTEMS OF THE ENGLISH TEACHERS AT MELLMAX HIGH SCHOOL

The Mellmax High School Grade Report Form

Because the grading systems used by any high school teachers must take into account the kind of report card the school uses, it is important to describe the report grading form used at Mellmax High School in some detail. (See Appendix 3).

The grade report issued every ten weeks at Mellmax High School consists of several fields which contain numbers representing one's, ten's, and hundred's. The teachers fill in the circles around the numbers in the appropriate fields: current period mark; current period exam; comment 1; comment 2; comment 3; and attendance. At the end of the school year, or at the end of the first semester, if the teacher is teaching a half-year

course, the fields on the right hand side of the form are used: semester exam, final exam, Regents exam, final average. The other fields (credits of exception, units, and RCT [Regents Competency Exam] or Prof [proficiency] exam) are completed by the computer program which scans the forms and prints the report cards.

Thus, the teachers at Mellmax High School find themselves ultimately adjusting their grading systems to arrive at figures and comments which will conform with what the report card allows and requires. For the five-week progress reports between report cards, the three comments must be taken from a list of 999 possible pre-scripted comments; many of the choices are redundant or innocuous, as the following list demonstrates.

039: "participates in class"

047: "pleasant"

074: "needs to study more"

058: "needs to pay attention"

122: "usually prepared"

While other comments are somewhat more specific, like #143: "struggles with material, but tries hard," or # 76: "student is rude and inattentive in class," the English teachers complained that finding the comment that says exactly what they want to say is next to impossible, in part because there is no clear pattern of organization among the comments, i.e., the comments are not categorized as "effort," "attitude," achievement," "behavior," and so on. (See Appendices 4.1, 4.2, and 4.3 for further examples.) Teachers are confronted with a three or four day window to fill in report forms for more than 120

students and come up with three comments from the list for each student that can capture what the teacher really means to say about the child.

Dealing with required comments

Many teachers resort to a "short list" of "safe comments" like "well-behaved" and "student is a pleasure to have in class" that actually tell little because they are open to interpretation and can mean what a parent would like them to mean. To avoid detection by the guidance counselors who check to see that three comments have been entered and may question the use of the same comments for everyone, some teachers a make a short list which includes redundant comments under different numbers. Thus, "pleasant," "a pleasure to have in class," and several synonymous phrases occur under more than a dozen numbers, allowing a teacher to quickly fill in the bubbles for perhaps three-quarters of his or her students with enough apparent variation to seem both conscientious and particular in choosing comments.

Other comments are tied directly to the numerical grades, so "student has a failing average," "there is a danger of failing this course," or "student's average is between 65-74" are obvious choices for filling in a third comment when a teacher finds himself at a loss for words or cannot locate the one appropriate remark he needs for a student whose numerical average is a seventy-one.

Harried teachers working to meet a deadline worry about provoking a parent conference with an ill-chosen comment—"Why did you say my child is 'very dependable' *and* has 'poor attitude?'"—or one that seems like a mismatch with the other two—"How can you say my child is 'putting forth excellent effort' and "is a pleasure to have in class' and then say she's 'in danger of failing the course'?"

A further fear is caused by the simple fact that the forms require some care in filling in the correct numbers. Since the comments range from 1-999, but the forms have places for ones, tens, and hundreds, a teacher may intend to fill in #832, "willing to help peers," but inadvertently fill in #842, "student has difficulty working with others." The reversal of intent means the teacher will have to explain the problem (if she can tell what happened in the first place). Had #632 been inadvertently filled in, "student needs to take instrument home to practice," it would be pretty clear a mistake was made—or the child is leaving his saxophone in the English teacher's room after school—but such errors can make the teacher seem careless or foolish, sometimes reinforcing an already negative image in the student's or parent's minds about the teacher. This can drive teachers to avoid using the more specific comments altogether, reducing the comments to irrelevancy and perhaps encouraging the teachers to adjust the numerical grade to better capture the message they hope to send.

Five-week report forms are usually distributed on a Friday, to be completed and turned in on the next Monday. A report must be prepared for every student, not just those at the extremes (i.e., those in need of a warning and those who merit special praise). Thus an entire weekend is spent under a kind of duress to "get the grades in." This may figure in the widespread feeling among the English teachers at Mellmax High School that the five-week reports have little value, or are not worth the trouble it takes to prepare them. At the end of the year, the first two comments are replaced with one comment for student attitude and one comment for student behavior. The end of the year directions (see Appendix 4.4) restrict the comments to the following:

D.) In the COMMENT 1 area, fill in circles 1, 2, 3,or 4 for student attitude.

     1 = SUPERIOR EFFORT SHOWN IN CLASS

     2 = EXCELLENT EFFORT SHOWN IN CLASS

     3 = SATISFACTORY EFFORT SHOWN IN CLASS

     4 = IMPROVEMENT NEEDED IN CLASSROOM EFFORT

E.) in the COMMENT 2 area, fill in circle 5, 6, 7,and 8 for student behavior.

     1 = SUPERIOR BEHAVIOR SHOWN IN CLASS

     2 = EXCELLENT BEHAVIOR SHOWN IN CLASS

     3 = SATISFACTORY BEHAVIOR SHOWN IN CLASS

     4 = IMPROVEMENT NEEDED IN CLASSROOM BEHAVIOR

F.) The COMMENT 3 area may be used, if you wish.

Thus the teachers find that their ability to express their evaluation of attitude and effort for the entire year is reduced to a very subjective summation that tells little about the details of the year's challenges, struggles, and achievements in the area of growth and development.

Dealing with reduced narrative grading

Among the English teachers at Mellmax High School, the general feeling was that the comments took up more time than they were worth, revealed little, and sometimes required an explanation beyond what they stated. The effort to find appropriate comments seemed disproportionate to their usefulness. This seemed especially true because a list of

three comments had to be sent to every student, not just those who seemed to require commentary.

The veterans, however, saw some value in the comments because they could serve as a kind of "red flag," giving some indication of why a grade might be as it was. The comments helped to "explain the grade" rather than assessing some other "non-achievement factor" like effort or behavior, despite the fact that they were often direct comments on such factors. Asked if the comments communicated what was intended, Hester Hypoint replied, "Oh, yes, definitely. Well, you know, either 'doesn't do homework consistently' or 'lost points by turning work in late' or 'does not complete outside assignments'. . . it takes me a long time to do those comments, a long time, because I go through and look at every kid's performance and how many homeworks they missed and how many tests they failed, and then base my comments on what I see."

"I don't know why we bother going to so much trouble, you know?" said Jeanette Jones, "What everyone looks at is the grade. As long as the grade seems OK, no one asks about the comments. Once in a while someone will ask about a comment, but usually only if the grade isn't what they like. Still, sometimes the kids ask. You know, they compare the comments. But they don't mean much." The other veteran teachers agreed, although both Karen Kistner and Mary Minton believed that even the rather bland comments that indicated a change in performance were worth using, since they sent a warning that a dip in the numerical average might not. As Mary Minton explained,

"Sometimes a child does poorly on a test and that pulls his average down just before we send out the reports. If you say, "Does poorly on tests," there's some indication of the

cause. But if you say, "is giving very little effort," then a parent might call to ask about that."

Karen Kistner viewed the comments as a mild way of letting students know their effort (or lack of it) had been noticed. "Kids will compare their report cards to see how they're doing compared with the other kids in the class. I try to find something positive to put on every report, even if the average is bad. But the grade is still the part they think is important—if the grade is bad, it kind of softens it to have a good comment, but if the grade is good, no one really pays much attention to the comments."

The veteran teachers had long been accustomed to the blue and white letters of the five-week progress reports. The blue letters, which indicated excellent performance, were sent to the high achievers; the white letters, indicating a deficiency, were sent to those with failing averages. While the text consisted of a simple statement that "your child is doing exceptionally well" or "your child is in danger of failing," the letter contained a checklist of possible reasons for deficient performance ("does not do homework; inconsistent effort; does not behave in class; needs to pay attention," etc.) and left considerable space for an individualized note to the parent about the child's performance. Hester Hypoint summed up the difference between the old system and the new: " With the blue and white letters, you only had to write to the kids at the top and the ones at the bottom. That gave you more time to say what you needed to say. Besides that, you were able to write what you wanted, and didn't have to go searching around for something close to what you meant.. . . No, the old system really was better. Now it's more work than it's worth." When asked whether the five-week progress reports elicited any response, virtually all the teachers in all three categories said no, at best only a

handful of parents contacted them about the reports. "In the 'old days,'" Hester went on, "parents were proud to get the blue letters—you know, they'd put 'em up on the refrigerator and the kids—and the parents—would compare the number of blue letters they got. But the new forms aren't like that. Nobody cares."

Among the novice teachers, the progress reports received a more neutral review. "I think they help a little. At least they give you a chance to send something home to the parents," said Betty Browne, "maybe they don't respond, but you know you sent them a warning." Catherine Carney, however, expressed this idea more negatively, "The five week reports are just a way for the administration to say, 'We warned you' if a kid fails. But that means *we* have to warn them. And sometimes a kid is just skimming along until the end, and then he fails. So you're not always sure what the best thing to put down on the report is." David Dutcher went further. "Those comments don't tell anything, really. It takes up so much time—for what? The average tells them if they're passing, and I've already told them if I don't like their attitude. And to send comments to everyone? That's just ridiculous. We have papers to grade."

The established teachers, to a person, concurred with David. Fiona Fisher, Frank Fender, and Gail Goodwin weighed out five week reports in the balance of usefulness vs. time required for completing the forms, and found them wanting. Frank Fender observed, "We use up a whole weekend filling in the little bubbles, and no one responds. Sure, a few kids feel good because you say you enjoy having them in class, but they already know that. And the ones who aren't trying already know that, too. You're not telling a kid with a bad attitude anything when you send that comment home—he already knows it, and usually his parents do, too."

Perhaps the differences in attitudes toward the comments reflected the relative ranks of the three groups of teachers. The novice teachers, being new to the scene, appreciated anything that let them feel they were establishing connections with the students and their parents. The veterans considered the letters a kind of safety net—years of experience had taught them that even if only one or two failures could be prevented by the reports, they were probably worth it. The veterans also remembered the previous system, and while the newer progress reports were a poor substitute in their eyes, they were better than nothing. But for the established group, most of whom consisted of young parents with children of their own still at home—unlike the veterans, whose children were grown, or the novices, who were only beginning their families—the loss of an entire weekend was being balanced against the apparent immediate usefulness, and the payoff was simply not sufficient to justify the time used in preparing the reports.

Grading Systems among the English Teachers at Mellmax High School

The pre-scripted comments, whether the limited set for the final report forms or the 999 for the five-week progress reports, made up a major part of the time requirement for preparing report cards. But they did not begin to approach in significance the one numerical entry on the card for each marking period: the grade. On the final report card, the four marking period grades and the final exam grade were averaged to get the final grade, and this number alone carried the greatest significance. The shared understanding of the weight of this number was clear: it meant success or failure in the eyes of the teachers, the students, and the students' parents.

This one number represented the summation of all that the teachers felt was important enough to be weighed into the grade, and the ways the teachers set up their systems for

calculating that one number varied, sometimes in interesting ways. While the grading systems devised and employed by the English teachers of Mellmax High School will be presented according to the three groups, some techniques are found in more than one group, and so the "common" approaches within each group are discussed first, and the exceptions to what seems typical of the group follow. Comparisons across and among groups are made where this seems relevant, especially when some other factor can be seen as explaining any overlap or difference in the groups.

Using a point system

Among the veteran English teachers, the "point system" prevailed as the dominant model for calculating the grades. Hester Hypoint described this system, widely used in American high schools. "I use a point system. Everything is worth a certain number of points, depending on what it is and how hard it is. So a quiz might be worth ten points, and a spelling test worth one hundred. A unit test might be two hundred points; essays are worth two hundred. Homework assignments might be worth twenty points. It depends on the difficulty. A big project might be worth three hundred points. At the end of the marking period, I add up all the points and divide by the total number of points they could have gotten if they got perfect scores on everything. That gives them their grade."

While the other veterans—with the notable exception of Jeanette Jones—all used a similar point system, their weighting reflected their understanding of the relative worth of the different assignments, so that the various weights were different, although similar, from one teacher's point system to another's. Projects in particular were subject to variation in weight: some teachers gave projects values as high as five hundred points,

making them the major part of a marking period's value, while others gave projects only two or three hundred points.

Philosophically, Hester Hypoint and Luann Lender were quite alike in their approach to weighting: essays weighed more than spelling or vocabulary tests, and quizzes on reading were worth only a few points. "If a student reads at a different pace, it can affect their performance on the quizzes, either because they are a little ahead or a little behind. So the test at the end of the novel carries the bulk of the grade for the marking period. The quizzes are just to keep them going," remarked Luann, "and homework assignments are usually little things that are for practice. So they don't weigh as much."

But Mary Minton and Karen Kistner differed on the value of homework: "I tell the kids they have to work to fail my class," said Karen, "if they keep up with the homework, they'll be all right. That weighs more than the tests." Karen felt that because many of her students came to her at an actual level well below the grade level, they needed the homework assignments to learn and practice. They were likely to do badly on tests, especially at the beginning of the year, and so she balanced her points so that the accumulated homework assignments could offset poor test scores. Hester and Luann frowned upon this practice, in part because the students were able to use their books, parents, or friends as sources of homework answers. Tests told the truth: who read and understood, who paid attention in class, and who was at the level they were supposed to be. Mary took the middle ground. Her homework counted significantly more than Hester's or Luann's, but not nearly as much as Karen's. "My students really have to do most of the homework and still do OK on their tests," she explained, "nothing should weigh enough that one test, say, can 'kill your grade,' because sometimes things go

wrong with a kid on a test. But they have to count, too, because they won't take them seriously if they don't count enough."

Jeanette Jones' point system was unique among all the teachers, not just the veterans. While every assignment netted points, all assignments were of equal value:

"I tell my students the first day, 'I'm not a numbers person. I don't play with numbers.' I don't double assignments, or triple them, or anything like that. Everything counts the same. . . . We do all kinds of assignments. A lot of reading, a lot of writing. I quiz them daily on their reading assignments. I've been doing a reading comprehension sheet every class with them. We're gonna do grammar. A lot of writing. And they're also taking the accelerated reader tests on the computer. I have all kinds of grades. . . . I just throw 'em in there. I don't double grades for major unit tests. I don't triple grades. I don't use a hundred points or, you know. Everything is just pretty much the same. Sometimes I'll throw out somebody's lowest grade—if you can see that they're usually performing at a certain place and one day they just blew it. I'm looking for patterns, I think, more overall of what their performance is."

Using categories

The other teachers also assigned points to various assignments; however, some categorized the assignments and awarded percentages of the total grade depending on the category. Over the course of the year, the novices who used categories adjusted them to reflect their experience using them, whereas the veterans were more likely to add or subtract assignments to adjust for experience. This suggests that the veterans were, as one might expect, more in tune with adjusting the pace of instruction based on performance as reflected in the grades their students were getting. The novices, on the

other hand, worried that their categories were creating lopsided outcomes in their grades, and so adjusted whole categories rather than the individual pieces within the categories. Betty Browne's efforts to find the balance were representative of the novice teachers' struggles:

"I have four categories: the writing workshop, the reading workshop, tests and quizzes, and homework. They get points for what they do in each area, and then I use the totals to get a grade for each part. Then I give twenty-five percent for each of the categories." Betty was finding that what seemed like a good system on paper might not be so easy to manage in reality. "I'm still working on finding the best way to do it. Especially the homework category. Because you think, you know, it's something they can all do, so it will help their grades. But then you have to be after them all the time. At first, I let them make up homework before the end of the marking period. But they didn't get things in, or they handed in all kinds of things the day before the grades were due. So I changed the rule to 'homework loses five points a day until it is a zero.' And pretty soon, they were all getting all zeroes for homework."

Alice Andrews created different categories. "I have an oral presentation every marking period, quizzes and tests, and projects." The weighting reflected the anticipated level of difficulty or the amount of time and attention the she expected would be required to properly prepare an assignment. "Quizzes count once, tests twice, and projects double or sometimes three times. I do grades based on effort." While grading and effort will be discussed in greater detail in the next chapter, this remark was apparently meant to indicate that Alice anticipated a certain level of effort would be required to complete particular assignments, and weighted them accordingly, not that she awarded points

based on perceived effort on the part of the student. Projects made up 60% of the grade, but were subdivided into parts that had to be completed according to a schedule; quizzes and tests made up 30%; and the oral presentation the remaining 10%. Alice, like Betty, penalized students five points a day for late assignments; for both Alice and Betty, these points were deducted from within the category, rather than from the average grade obtained for the report card.

Catherine Carney, also a novice, had only two categories: 60% writing, and 40% quizzes, tests, and handouts (i.e., daily seatwork assignments). The second category covered "anything that doesn't fall under writing goes into my 'quizzes and handouts' category, and that's 40%. It doesn't matter whether it's homework or not, or in class or not." The writing assignments were divided into halves, with 50% of the grade being based on content, and 50% of the grade based on format (i.e., structure, mechanics, grammar, and so on).

None of the veteran teachers used categories to contain—or ensure—the effects of certain classes of assignments on grades. If a unit required many homework assignments, each was given its points, and its total value became part of the divisor in the equation for the grade. The novices who used categories of assignments believed that it gave better balance, so that one kind of assignment did not present a problem if the ratio of assignments changed from marking period to marking period. Thus, if there were fifteen homework assignments in the first marking period and only ten in the second, homework still counted for a fixed percentage of the grade. The use of categories also helped them to be sure they included a variety of assignments in their gradebooks; as Betty Browne put it, "Some kids are better at some kinds of assignments than at others. This gives them

some points for their strong suit every marking period." Looking at it from another perspective, however, one could say that it also meant most students probably lost points on their "weak suit" every marking period. This practice also reflected the fact that the novices generally gave fewer assignments over the course of a marking period, having perhaps ten or twelve grades to average in a ten week marking period, whereas the veterans often had thirty-five or forty different assignments of a wide variety in a ten week period. Veterans, then, seldom worried that they hadn't "covered the bases," and probably felt that the use of categories would complicate their calculations precisely because they had so many more scores to average.

Combining points and categories

Depending on the units covered in the marking period, the established teachers—and some of the novices—used points or categories or hybrid systems that used points in one marking period and categories in another. If a period had a "major project," the project might become a category of its own. Frank Fender explained, "When my students do the research project, it makes up a set percentage of the grade for the marking period. If they don't do it, they lose all that part of their grade." If the system of points were used alone, then students could conceivably offset the loss of points on a project they decided they didn't want to do by accumulating points with high scores on things they found easier. By making a major assignment into a self-contained category, students would be obliged to do the assignment. For example, if the assignment weighed 30% of the grade for the marking period, the highest average that could be obtained without doing the project would be 70%, leaving only a five point margin over failing. High scores—even perfect ones—on all other assignments could still net no more than a grade of seventy.

The teachers using point systems without making major projects into categories would have to weigh such projects very heavily to "force" the student to complete them; the danger here is that they weigh so heavily that weaker students can hurt themselves with low performance because the effect isn't limited to a set percentage of the total grade. Thus the strategy of adjusting the grading system to reflect the importance of assignments during a particular marking period is a useful one.

Taking Non-Achievement Factors into Account

Although the inclusion of the comments on the report forms and progress reports seems to facilitate the separation of achievement and non-achievement factors in grading, some of the teachers felt that certain non-achievement factors like attendance, attitude, effort, and participation have such significant effect on achievement outcomes that they incorporated these directly into their grading systems—although perhaps not for every student.

As Catherine Carney explained, "I do take my special ed students into consideration differently than I do my regular ed students because they have IEP's—spelling, certain things I can't hold them to; test modifications, whatnot, . . . so I tend to look at effort as —'Am I teaching them a work ethic?' I guess is what I'm thinking about. . . .If they're not giving any effort to it and they have all these things, I'm going to hold them to it; I'm going to nail them. But if they *do* have a work ethic and they're trying pretty hard, they just don't have the raw material to do it, I know I'm easier on them, I know I am."

Ellen Enders described the use of a participation grade in group projects. "Now with the block [schedule], we've been doing more participation, and those things work great for the kids who may not do so well on the test because maybe they're great artists or

they're great at other things, so that they're on task, that they're participating, that they all have an active role and I assign what those roles will be. …so they're all held accountable and they all get to do something they like."

Asked about how she assessed satisfactory or unsatisfactory behavior, Jeanette Jones, answered, "Well, I don't give a grade for behavior in my classes. I guess I stopped doing that years ago." But when asked whether that affected the comments she sent home, she replied, "Yeah, oh, sure, and it also might affect, you know, if you said, 'Well, this person is trying so hard and has a sixty-three, you might not give them a sixty-three. But if somebody really goes out of his way to give you a hard time, you might *not* give them a sixty-five." Asked if the kind of behavior she observed on the part of a student influenced her grading in any way, she answered, "Well, yes, it can. You have to be honest (laughs). I like to think I'm totally objective, but there are times when they've crossed the line, and you're not going to bend, either."

Flexibility and Fairness

While most of the teachers reasoned that non-achievement factors influenced student performance, they did not assign values to the factors in any direct way, so that points were not awarded for good conduct or effort or deducted for bad behavior or poor effort. Yet many, like Jeanette Jones, admitted that when a student was "on the bubble" between passing and failing, they might "go back and find points" if the child had tried very hard or was well behaved on a consistent basis. No one said that they would re-examine their calculations to see if points could be deducted from a student whose behavior had been especially galling, but several admitted that if such a student were "on the line," they would not make an effort to "find" points for him.

The teachers did not view this practice as unfair; it was within their rights to exercise this power, and those who had cooperated could be awarded a "little boost," as Catherine Carney put it, whereas those who had not had no right to expect it. Being left where they were was what they deserved in return for their non-compliant attitude.

In the first interviews held at the beginning of the school year, all the teachers were asked if they were fair in their grading, and to a person, each replied—usually with little or no hesitation, "Yes, I am." Asked the same question again at the end of the school year, with final grade reports on the horizon (and after several interviews exploring their grading practices), virtually every one of the teachers hesitated and replied, "I try to be," or words to that effect. Mary Minton, the most senior of all the teachers, replied, "I think I am; I'm just." Much like a judge at the sentencing of a criminal, each teacher felt they had legitimate authority to bend their own rules—or not—as the circumstances of the individual before the bench gave warrant, provided that such slight accommodations were in keeping with their concept of what was fair.

Dealing with Subjectivity: Rubrics, Checklists, and Commentary

The English teachers in this study were comfortable with subjectivity. When asked about the balance between subjective and objective assignments in their courses, all remarked on the subjective qualities of writing assignments, and estimated the balance at either "50/50" or "60/40," with the larger part being subjective assignments—i.e., assignments requiring interpretation and personal judgment in order to be graded. Because the "exit exam" for high school English in New York State is comprised of four essays, all the teachers felt that it was part of their responsibility to prepare their students to write essays of the kind that would be required on the exam. Ironically, the English

standards for NYS include "reading, writing, listening, and speaking," but the test so heavily samples writing that the teachers saw reading, listening, and speaking as holding a distant second place to writing. While the exam includes a few short reading passages and some multiple choice questions about them, the teachers felt that the level of ability actually required to do well on the reading passages was elementary, and so used their reading assignments chiefly as vehicles for deriving essay assignments.

While one might expect that grading would involve judging the actual processes of writing, writing done in class was not assessed until after it had produced a product. Thus, writing assignments often required first draft, editing, and final versions, but each of these yielded a completed product that was graded as such. The first draft often weighed more than the final product. "I weigh the first draft more because I grade it before they do the revisions. Basically, I tell them what to do to improve it, so if they do what I tell them, *they're* only following directions. So the final draft doesn't count as much as the first," explained Frank Fender. For the 12th grade research paper, Alice Andrews weighed the first draft three times, and the final version two times; Catherine Carney's 12th grade research paper's draft counted twice as much as the final version, because the students "are only making specified changes to the first draft."

All of the teachers used rubrics or checklists when grading essays. While a checklist usually identifies required elements, and a rubric develops this into a set of criteria for various levels of quality within those elements, most of the teachers treated the two as synonymous, sometimes actually using a rubric, but other times using only a checklist that identified those pieces of the assignment that the student had to include. Some of the rubrics used by the teachers are included in Appendices 5.1, 5.2, 5.3, and 5.4; the wide

variety and differing levels of sophistication suggest that for some teachers, the rubric merely provides a reminder of what is expected, rather than a grading guide, leaving them to make their judgments according to their established practice, but ensuring that they "give credit where credit is due."

The most involved of the "rubrics" was one used by David Dutcher, who, after transferring to another school district, found that his rubric continued to be just as useful. The rubric divided the essays into parts and assigned values to each of these parts: 60% went to "organization, content, and development"; 30% went to "mechanics and conventions—punctuation, spelling, grammar, sentence structure, even legibility"; the remaining 10% went to "style and wording, vocabulary and usage."   Within these categories, however, a great deal of judgment had to be applied, and, while his students were made aware of the rubric's details and how David would be using it, he still found that students disputed their scores.

"I had a girl who came to me with an essay I had returned. She didn't like the grade, and I took some umbrage at her remarks. She told me 'English is a matter of opinion,' and if she took the paper to Miss Grimm (another English teacher in the school), she would have gotten a different grade on it. And she's probably right. That's the real problem with English—everyone has a different idea about what good writing is." Still, he was sure that once students understood how a teacher graded, they could interpret the grades they received in keeping with the message the teacher was actually trying to send.

The other teachers made similar remarks about the ways that grades reflected the teacher's judgment, but were not uncomfortable with making such judgments.   All expressed doubts that there exists some sort of Platonic "ideal essay" of which the essays

they received were poor imitations, and allowed that it was unlikely that any amount of discussion among English teachers and grading experts would ever arrive at a method of grading so certain that one could declare with certainty, "***This* is** an 85."

Because of this shared awareness of the limited usefulness of the numerical grade, all of the teachers felt obliged to add commentary to the papers they graded. This feedback sometimes explained the nature of an error, or offered a better word, or suggested an improvement. Often it "defended" the grade, and remarked on the quality of the essay in relation to the rubric. Or it could address the content of the essay, remarking on its interest despite the low grade based on organization or mechanics. In other cases, it offered encouragement despite the evident weaknesses displayed in the essay, or gave praise for work that showed improvement.

Gail Goodwin described the likely distribution of comments on essays and research papers. "The papers that get the lower grades get a lot of comments on them. Those are the students who need more explanation of why they got the grade the got. They also need more encouragement. I always look for something positive to say on a paper, no matter how bad it is. If all they get is a low grade and a lot of negative feedback, why should they keep trying? But the better papers don't get so many comments—maybe 'good job' or 'keep up the good work.' The grade speaks for itself."

While the other teachers echoed Gail's sentiments, no one explained why the higher numerical grades were somehow self-evident, and the lower ones required commentary to support them. Perhaps it was because students were content with the higher grades and saw no reason to seek an explanation for them—or perhaps the grade was sufficient encouragement in and of itself.

Like the comments on the grade report forms, the purpose of narrative grading was reduced at Mellmax High School to fulfilling a kind of support role for the numerical grade, either by providing explanation, adding detail, giving a warning, or suggesting some strength in another area not displayed through the assignment to which the comments were added. Properly understood, narrative grading means that the narrative is the grade; however, the nature of the school-wide grade report at Mellmax High School meant that narrative grading never stood alone, and served to augment numerical grades.

Second Chances

While each of the grading systems the teachers had devised could be used to account for every graded assignment, sometimes students handed in work that was not what was expected, either because the assignment itself was unclear or because the student did not perform at his or her usual level. Almost every teacher allowed for "second chances" in some area or another, with only a few insisting that an assignment, once completed and handed in, was indelibly etched into the gradebook.

Hester Hypoint allowed students to retake vocabulary tests. "That's one of those things that you either learned or you didn't. So if a student fails a spelling test, I will let them retake it—*after* they study for it. Since they have a spelling test every week, it is a significant part of their grade. Besides, for the weaker students, spelling is something they can do. So letting them do them again helps their grade. And you can either spell the words now, or you can't." The new grade replaced the old grade, so a student who would study for the test could indeed improve his or her average by retaking vocabulary tests.

Other teachers—most, in fact—allowed students to re-write essays if they were unhappy with their grades. But not every teacher replaced the old grade with the new one. "I give them the average of the two grades," reported Fiona Fisher, "because often the reason the first essay was no good was because they did it in a hurry or didn't really put any effort into it. If they're going to re-write an essay, they have to come in and conference with me. We talk about the first essay and I make suggestions about how to improve it. Then they can turn in a new version. But that means I did some of the work for them, and that's another reason I give them the average of the two grades." Luanne Lender agreed, but found that allowing students to conference and then rewrite essays was "a great learning experience for them. Often they write essays that are really good after they talk to you about them. I've had essays that were just abysmal—you know, in the thirties or forties, and the kids would rewrite them and they'd get nineties." She tempered her enthusiasm, however, with the bitter observation that, "most kids won't even bother to try if you give them a second chance."

Teachers did not allow second chances on unit tests or most quizzes, however; because the answers were usually revealed shortly after the test had been given, students had to keep the grades they had received. And research papers, which usually required a first draft and a finished version, already had a kind of "second chance" built in—although the heavier weighting of the first draft discussed earlier meant that students who learned a little later how it should have been done suffered more than those who turned in a strong first draft.

Policies regarding homework assignments varied; some teachers accepted late homework, some deducted a set number of points for each day the assignment was late,

some allowed students to hand in back homework any time before the marking period ended. Still others demanded that homework be handed in on time, or it received an automatic zero. Often this reflected personal experience with the problems a steady stream of homework presented: piles of late homework to be checked on the eve of report card submission day could be a problem if a teacher were too lenient. If a teacher were too strict, however, a large number of zeroes for homework could mean a disproportionate number of failures in one section or another.

Betty Browne mused over her difficulties in devising an effective policy on late homework. "At first, I let them turn it in any time before report cards. But then I'd get like hundreds of papers at the last minute. I didn't give grades on homework anyway, but they got credit for doing it completely. So I'd have to go through and check them all. And a lot of kids wouldn't do it anyway, so it was discouraging. Then I told them they had to have it in before the end of the week it was due. That didn't work. I tried no homework, and they fell behind. So then I tried a 'zero tolerance' policy—the homework had to be in the day it was due, or it was a zero. But I had so many failures at the end of the first marking period, I had to change that."

The guiding principle in granting second chances seemed to be benefit to the student without detriment to the teacher. Another clear reason for giving second chances was the need for grades to conform to the tolerances of the larger school system's grade report system.

Wiggle Room

Giving second chances on certain kinds of assignments was one device for providing a kind of "wiggle room" in a grading system—wiggle room for both the students and the

teachers. Other ways for giving some room for movement of grades became apparent when the teachers were asked, "What do you do with an assignment that's a disappointment to you?" Some of the teachers interpreted the question as referring to an assignment they had given that was badly designed or explained so that the class as a whole did poorly. "Oh, I throw it out," said Ellen Enders, "I won't count an assignment where the bad grades are really my fault."

"I look very carefully at the assignment and the papers to see if it was my fault," remarked Jeanette Jones, "if it was a lousy assignment, I won't count it. But if I think it was a good assignment and they just didn't try, I keep it. If it's just a few students who misunderstood, I conference with them and see if we can make it clear to them what was expected. And I let them do it over."

Luanne Lender was a bit harsher. "I throw it in the trash. Right in front of them. I've used these assignments for years, so I know they can do them. If a class does really badly as a group, I tell them, 'I've never seen papers this bad,' and I throw them out. Sometimes I won't really count them, but sometimes I will. It depends on how they react. If they'll come and ask for help with it, or promise to do a better job, I may give them a second chance."

Discussing individuals' papers that were disappointing—not living up to expectations, or, in Jeanette's words, "not part of their pattern of performance,"—most teachers were willing to talk it over and allow a second chance, even if this were not the policy for that kind of assignment. Nearly all the teachers admitted that if second chances were offered, they were more likely to prompt the students they perceived as cooperative to re-do the assignment. While all were scrupulous about informing the

entire class if they would accept a re-written assignment, most of the teachers admitted they were less likely to pursue the less cooperative students after they had fulfilled their moral obligation to inform them of the option. Many of the teachers were willing to "drop the lowest grade" either in a given category or across the board; for some, this was a stated practice in their grading policy. Almost as if to say, "We all come up short sometimes," dropping the lowest grade seemed built in to some teachers' systems to help build their students' trust in their teacher's sympathy for them.

Another device to allow for some adjustment of grades was the use of letter grades for assignments, like essays, that required subjective judgment. While all the teachers used points in order to calculate averages, most used letter grades on essays because numbers seemed too precise. The letter grades represented a range of possible scores (e.g., A = 92-100; B = 81-91, and so on). Plusses and minuses narrowed the range for each letter grade. The letters could be translated into numbers using the possible ranges, and this allowed for a built in flexibility in this area of the grades when it came time to determine averages for reporting. While some teachers automatically converted letter grades into the midpoint value for the range (so A = 96; B = 86, and so on), others could choose a number within the range, giving them "points to play with." The decision might be based on a perceived pattern of overall performance, or it might reflect some subtle judgment regarding a non-achievement factor like attitude or perceived effort. It could also reflect a positive or negative feeling about the student due to classroom conduct; teachers were generally vague about how such decisions were made.

David Dutcher, however, used no letter grades—his rubric for grading essays translated all parts into points, and he could indeed explain how a 94 differed from a

95—at least, according to the points awarded in each category (things like spelling, mechanics, sentence structure, style, etc.) and according to his own judgment, which he acknowledged as part of the system. Within a category, however, justifying 9 points for "style" versus 8 points was somewhat harder. Here, he, like the other teachers, admitted that some of the values were reflective of a subjective, and perhaps personal, judgment about, or even reaction to, the work.

While this lack of wiggle room could make David's grading seem more certain, and therefore more efficient and less time-consuming, no one among the teachers worried more about the challenge of grading. "I spend *days* grading their essays; sometimes I re-read them all to be sure I'm doing this right. I don't think anyone spends more time on grading than I do." Quite the opposite, the use of letters by the other teachers seemed to facilitate their grading by eliminating the need to agonize over the difference between a 74 and a 75. And when it came time to convert the grades into numbers, anything within the range seemed fair—in part, perhaps, because the actual paper could not be called to memory in its details.

One of the interesting ways in which "wiggle room" could be built in to a system was explained by Betty Browne. "In the writing workshop, I use a rubric to grade. You know, like the one for the [New York State] Regents exam. I grade the papers from one to six. But sometimes I make a little note, so a paper could be a 'high five' or a 'low five.'" The creation of a range within the narrow limits of the rubric allowed Betty to move some grades up and others down (in small steps) at her discretion.

Fitting it all to the Larger System

Features of Mellmax High School's grade reporting system clearly influenced the teachers' grading systems:

- The numerical grades are clearly assigned greater importance than the comments on both the progress reports and the final grade reports, often causing the teachers to use the comments as "explanations" for the numerical grades;

- Comments are reduced to pre-scripted choices which may not clearly express the teacher's intent, giving rise to frustration in using the comments and encouraging the use of a "short list" of "generic" comments;

- Promotion or retention decisions are based on the numerical grades alone, so that teachers may give weaker students who are perceived as doing the best they can a slightly higher average than they actually have. Conversely, teachers will not extend this help to students they perceive as uncooperative or hostile, leaving their averages as calculated;

- There is an unstated but understood "acceptable rate" for failures; while no set number is ever actually given, all teachers have a sense of what the system will accept for their grade level and the time of year (fewer failures being tolerated on the final report form).

- Like the implicit rule that failures should be kept to a minimum, the district employed an unwritten policy that in the first marking period, no grade lower than a fifty could be given. While the novice teachers accepted this as a rule, the veterans and established teachers observed it at their discretion, refusing to automatically give a fifty to anyone they perceived as lazy, vexatious, or utterly incapable.

- A regular timeline for submission of grades means that the grades are not good representations of the student's actual performance; for example, if a new unit has just been begun, the student's success in the current lessons may be overshadowed by his previous shortcomings in the last unit. This works in the opposite direction as well. Because of this, teachers may hesitate to "count everything as it is," knowing that the report card grades are "fixed," and if the average is calculated today rather than next week, it will be detrimental to the final average.

All of these features can affect both the devising and implementation of the teacher's grading system. Usually the more experienced teachers build some flexibility into their systems—as in allowing second chances or using letter grades for some assignments so that they have some room in converting the grades into numbers. But all the teachers sometimes find themselves making adjustments due to the demands of the larger system.

Novice teachers are more willing to adjust their grading practices to obtain results they believe are desired by the school district. That is, as the end of the marking period approaches, the novice teachers begin looking at the obtained averages based on their stated grading system. If they find that it seems that there will be "too many failures," the novice teachers begin making adjustments. These may be varied according to individual students' needs for more points, or they may be wholesale changes in policy.

Alice Andrews, as yet untenured, remarked that, "When I get to the end of the marking period, I take a look at the averages to see how many students are failing. If I have a lot of them and some are on the line between passing and failing, I'll take another look at them. I might not count something, or I might not average in some of the zeroes

for missing homework. Things like that. But I never give them something that changed their grade a lot. But you know, you can't have too many failures."

Betty Browne, also untenured, adjusted her grades in a different way to limit the number of failures. "At first, I told them homework had to be in on time or it was a zero. Pretty soon, everybody was failing because they weren't handing in their homework. Well, not everyone, but a lot of them. So I had to change the weight for homework. I made it count less."

The last of the untenured teachers, Catherine Carney, resented the fact that "you can only fail so many people. But, you know, they'll say you're a bad teacher if you fail more people than the other teachers. And if you're not tenured, you know? So, yeah, I'll take out some assignments, or give the people who are failing a chance to do extra work to make up for what they haven't done."

The experienced teachers were less willing to make changes in their grading systems in order to adjust the outcomes and reduce the numbers of failures. Because of their longer experience (and tenured positions), they had settled into grading systems that were designed to keep failures contained to levels they believed were fair. They seemed to be largely comfortable with their grading schema, but sometimes found themselves in situations that required change or compromise. Sometimes this was the result of an error on their part, and, having learned from this mistake, they adjusted their policies to prevent the situation from occurring again.

As a case in point, David Dutcher, whose faith in carefully planned grading rubrics seemed unshakable, found himself in a new assignment after changing school districts. Here he was expected to assign an independent reading project that involved writing a

book report and assembling a list of "new" vocabulary words. Students chose their own books to read, obtained David's approval, and then wrote up their book report according to David's carefully designed rubric. In order to encourage the students to develop larger vocabularies, David added as the last element in the new rubric a provision that students would receive one point for each new word they listed, defined, and used correctly in a sentence.

To David's surprise, an especially poor book report was handed in "by one of my better students. This kid is smart, but he's kind of a wise guy—always testing his teachers and that kind of thing. Well, I grade his report using the rubric, and he gets a 73. But then I find a list of words at the end—forty-five of them!" Now, according to David's own rule, the lad gets 45 points for his increased vocabulary, and that gives him a grade of 118—and it counts *double*. After an argument with the boy, a call to his parents, a call from his parents to the principal, and a conference with the principal, David was told to back down, give the points, and change the rubric for next time.

David's case is unusual for an experienced teacher. It demonstrates that teachers sometimes make mistakes in their thinking about grading—and can be very stubborn about admitting them or correcting them.

Frank Fender took a more philosophical view of administrative intervention in grading. "I have the 'average' seniors. By this time [the end of the year], they are not working anymore. You can't just fail 'em. So for the seniors, homework is optional. If you do it, you'll get credit, but it won't count against you if you don't. But you still have to pass the tests and do your papers. For the 10th graders, though, it's different. They

have to take the Regents exam next year, so they have to be ready, so I count their homework because they need to really work this year."

The veteran teachers were at ease with their own systems. Undoubtedly this was in large part due to their long experience and the development of techniques for dealing with complaints or charges of unfair treatment. Pressure from the administration was met with a grade book and a comparison with the previous classes of many years. Having taught the parents of the children they were teaching now, they were much more at ease when dealing with parental complaints—or when deflecting pressure from the school district.

Ironically—or perhaps intentionally—the veteran teachers in Mellmax High did not want the role of final gate-keepers: none of the veterans, despite their years of seniority and long experience, taught twelfth grade English, where resistance to district pressure to "push people through" would be met with the greatest dissatisfaction. Hester Hypoint revealed that it was not uncommon for the principal to visit her at the end of the school year to pointedly inquire about her larger-than-the other-teachers'-failure rate.

"I get out my grade book, and I go over the assignments and the grades these kids have gotten. I say, 'Look, here are their averages.' If you don't do your work, you fail. That's just the way it is." Asked if the principal insisted that she change any grades, she replied, "No, but you never know what they do after the year ends. And there's always summer school. I never know for sure how some of these kids I fail make it into tenth grade in the fall. I think if I taught the seniors there might be more pressure, you know, to kind of push some of them through, but there isn't much of that with the ninth grade."

GRADING SYSTEMS AND THE MORAL

We have looked at the grading systems the teachers at Mellmax High School have established for determining their students' grades. The many features of these systems and the way the teachers actually apply them are driven by three considerations: doing what is fair (or just, as Mary Minton says); keeping the students moving forward; and fitting the process of grading to the product expected by the school district. The theoretical framework established in chapter one can now be used to consider the moral implications of the grading systems examined above.

Truthfulness and Grading Systems

Jackson, Boostrom, and Hansen (1992) posit the assumption of truthfulness as a basic moral element in instruction. While observing that teachers feel compelled to consciously enforce this assumption during testing by minding their students more closely and in an obvious way, truthfulness pervades not only the actual testing conditions—where the students' truthfulness must be ensured—but also the entire creation of the test, its administration, and its assessment.

If we accept Karmel's broad definition of "test," virtually all assignments, even the ungraded ones, are tests. That is, everything a student is asked to do contributes to the informal and formal judgment the teacher will make upon him or her. The importance of truthfulness, then, becomes obvious. As Jackson et al observe, the students must give honest answers to the best of their ability so that the teacher can draw accurate conclusions about their level of understanding, ability, or mastery. Teachers must assume truthfulness in this way on the part of the students. In return, however, the teachers must act truthfully. Tests have to be graded accurately, and should be

interpreted truthfully by the teachers. The grading systems devised by the teachers at Mellmax High are largely based on points, and the implications of using points to reveal the truth about a student's performance are significant.

The chief advantage of a point system is the ease of calculation it allows. Once weights have been assigned to the various kinds of work (if they are), calculating the average is a simple matter of summing the points earned and dividing by the number of points possible. This means that the average can be calculated at any point in the course of the marking period using only the assignments completed up to that point. If a teacher gets ahead or behind in her schedule, it makes no difference; the calculation can be performed at any point. The moral question here is whether or not the average calculated at any given moment—especially if it is in the midst of an as yet unfinished unit—is actually revealing the truth about a student.

Many of the teachers at Mellmax High were aware that this was a problem. Especially at the beginning of the year, when they were just getting started, they worried that the first five-week progress report was not actually truthful. For one thing, because of the school calendar, actual instruction had only been taking place for three weeks, and teachers had collected little more than initial information about their students. Some students start slower and some start faster; the initial "read" a teacher gets on a student can be false, and sending a report so early seemed to most of the teachers worse than premature—it was wrong, because they could not yet know the truth because it was not yet clear to them.

Precisely because the point system can take an average at any time, and the teachers felt they "had to have a grade," (indeed, the district required it) many teachers hastily put

together a few assignments, graded them, averaged them, and announced to their students, "These are your averages at the moment; they can change with the next assignment. So don't be discouraged if you're not doing well. We have lots of time left in the year. But don't be getting overconfident either. Just one zero will change your average to failing." This is the dilemma of truth in grading when averages can be calculated at any time rather than at the actual conclusion of things, whatever those things—units, lessons, school years—may be.

Because the teacher controls the weighting of assignments, there are ways to manipulate the weighting to the advantage (or disadvantage) of the students. And this creates a new threat to truth in grading: difficult assignments can be downplayed if students do poorly, and easy ones can be weighted more heavily than they merit. A teacher desperate to limit the number of failures in a class can alter the averages by padding simple assignments with points. Dropping the lowest grade can be an inflationary device to keep class averages from seeming too low. While the grade book can serve to prove the truthfulness of the calculations, it cannot prove the truthfulness of the weighting and distribution of points.

Jeanette's search for patterns in performance seems intuitively appropriate, but presents a problem—at first, a change in performance which may signal a new pattern will be seen as a fluke if it is a drop in achievement, and then Jeanette is apt to throw it out. By making a judgment that the grade doesn't fit, she may not be telling the truth about the child's performance—and may be preventing his or her parents from realizing something is beginning to go wrong.

But Jeanette's equal weighting system has one advantage tied to her willingness to drop the lowest grade if it is not typical of a pattern of performance. While dropping the lowest grade in another teacher's system can pose a problem because the various weights mean one student may be gaining more than another—say, by dropping a 200 point test instead of a twenty point homework assignment—Jeanette's assignment of equal value to all work means that this is not an issue. Dropping the lowest grade on one assignment or another makes no difference beyond the extraction of the one grade.

The moral question that arises, however, is whether it is fair to drop the lowest grade only when it fails to fit in a pattern—or, for that matter, whether it is fair to drop it at all. If one student has a pattern of consistent performance, dropping the lowest grade may raise his average, but probably less than it would for the student whose performance on the lowest grade is inconsistent with past performance. Hence, the student whose performance is more consistent does not gain anything for his consistency of achievement, whether high or low, whereas the inconsistent performer improves his overall average by not being held to account for the atypical poor performance.

Using categories of assignments can be seen as a way to help guarantee truthfulness in grading if English is understood as a complex combination of skills, knowledge, growth, and achievement. A teacher who uses categories is attempting to guarantee that all the bases are covered, assessed, and incorporated into the grades she gives. Yet categories suffer from the same potential for manipulation, and an imbalance among the categories can prevent the student or his parents from knowing the truth about where he stands in the total English/Language arts picture. For example, if homework is a category, but many homework assignments are given in one marking period and few in

another, the grade calculated using the category approach may not be true; a student may have done all thirty homework assignments in the first marking period, and accrued her fifteen percent for them, but missed one of the three homework assignments in the next period, and lost five percent for that.

Another issue involving truth and the calculation of averages at fixed times in the year arises from the fact that, for most students, grades obtained at each marking period are "frozen." The fifty-eight received in the second marking period—the moment the average was calculated—becomes a fixed value to be averaged with the other marking period averages and the final exam grade to determine the "final grade." This use of fixed grades taken at specified moments in time can distort the truth; if the student comes into class at the beginning of the year far behind in ability, his early grades will reflect this. If he advances, perhaps remarkably, beyond his entry level, this will be reflected in his later grades alone. What he gets at the end is the average of all these fixed grades— which certainly do not tell the truth about him as he is at the end of the year. The underlying question here is whether the final grade is meant to capture the truth about the year's work taken as a whole, or to capture the truth about where the student is by the time she has reached the end of the course.

A more complex moral issue regarding truth arises from the fact that teachers *want* their grades to tell the truth, and are by no means unaware that points and categories, while serving an important purpose, have limitations that cause them to distort the truth about their students. Efforts to counter distortions caused by the limitations discussed above can include the same steps that can create the distortions: throwing out "bad" assignments; discounting the lowest grade; shifting the balance of points or categories;

adding bonus points or deducting penalties; substituting one assignment for another. The guiding principle for the scrupulous teacher is the desire to reveal the truth.

For some teachers, the grade is the *means for determining the truth*, and for others it is the *vehicle for telling the truth*. For the first group, this means that grading systems must be painstakingly designed and carefully balanced to accurately capture the truth about their students. For the second group, it means that their system must be constantly adjusted to prevent its built-in mechanisms from distorting the truth.

The "truth finder" judges non-achievement factors like effort, attitude, and interest by inference from the grades, which have revealed the truth. The "truth teller," on the other hand, uses informal observational assessments of effort, attitude, and interest to adjust the system to arrive at grades that reflect this information. The moral complexity of truth in grading is further increased by the fact that all conscientious teachers are *both* truth finders and truth tellers.

Worthwhileness and Grading Systems

Worthwhileness, the second of the substructural elements of the framework from *The Moral Life of Schools*, rests on the idea that schools are places where one is aided in developing skills and increasing knowledge. To repeat Jackson et al's insight, "This implies that the people in charge care about the welfare of those they serve and only ask them to do things that are expected to do them good" (p. 25). Where grading systems are concerned, worthwhileness is a serious moral obligation—the grading process itself must be worthwhile, and the assignments being given must be worthwhile.

All of the teachers at Mellmax High believed their assignments were worthwhile, at least in a general sense. Some students, they acknowledged, would not find every lesson

worthwhile. The use of heterogeneous grouping meant a wide variety, even disparity, of talent in almost every class, and the teachers knew that for some students some assignments were too easy, and for others, too hard, and that in both cases the worthwhileness of the assignments was, therefore, low. Yet even when this was so, students were expected to do the assignments for other reasons—for practice, for the sake of making the attempt, even for developing patience. This was not "doing them good" in an academic sense, but it was in a moral sense, since it helped to develop their character. And thus some level of worthwhileness could be assigned to every assignment, even when it seemed to lack worthwhileness on the face of it.

More significantly, if the assignments being given are expected to do the students good, then the grades they receive should also do them good, since the grades should reflect the good intended for the student. This does not mean that the amount of good done is indicated by the level of the grade received, but that the grade serves to indicate how much has actually been learned, and makes this clear to the student receiving the grade. A low grade on a difficult assignment should not be understood as a "bad grade," but an indication that there is more to be learned. But learning what remains to be learned requires that it be worthwhile to do so, and this is where grading as a moral process can suffer.

Because grading is often treated as a process of fixed-moment reporting, it becomes tempting to take a grade and then move on—and away from the material not yet mastered. "We're finished with the unit. You have your grades. Now we begin the next unit." The message sent by remarks like these is that what had been presented as worthwhile enough to struggle with at the start of the unit may not be worthwhile to

attempt to master now that the unit has ended. A passing but undistinguished grade, in particular, seems to suggest that the worthwhileness of the assignment is actually somewhat limited, since the unlearned portion can be safely left behind. Indeed, the fact that assignments from an earlier time could not be re-visited later suggests that the worthwhileness of such things is time-sensitive; having passed the time when it "counted," students were expected to leave it behind, undone and unlearned.

All the grading systems used at Mellmax High School were based on the assumption that once a unit was completed, the students had learned all that they could. But this is not the same thing as learning all that one *should,* and this is where worthwhileness is undermined by reporting systems that do not allow for mastery except for those who are quick enough to learn the entire lesson in only the time allotted. The teachers were aware that this happened sometimes, however, and some would allow students who had not mastered a lesson to learn the material at their own pace, return to take the test again, or rewrite the paper, and replace the first grade with the second. Hester Hypoint's re-takes on spelling tests proved to her students that vocabulary was worthwhile. Luanne Lender's detailed plan for re-writing an essay, the grade on which had disappointed her or her student, also indicated that such assignments were worthwhile. Refusal to allow students to re-do assignments that they had simply not prepared for was not a rejection of the assignment's worthwhileness, but a practical consideration: students who acknowledged an assignment as worthwhile but failed to do it did not deserve a second chance, and the time required on the teacher's part to allow such students to work as they pleased rather than as they should was simply not available.

A final observation about the moral implications of grading and worthwhileness comes from a consideration of the practice of "throwing out" certain grades as the end of the marking period approaches. If all assignments are indeed worthwhile, as every teacher attested, how is possible that some are sufficiently less worthwhile to be discarded because the marks are low? If the same assignment is thrown out for all students, this suggests that the assignment may not actually have been worthwhile in the first place. If the "lowest grade" for each student is being thrown out, this could be any of the assignments given in the marking period, suggesting that all the assignments are equally *not* worthwhile. Jeanette Jones' policy of giving equal weight for all assignments distorts the worthwhileness of particular assignments by equating all work, whether simple or demanding.

However, if one considers all the purposes of grading—one of which, after all, is to advance the students and to limit failures, then worthwhileness may be partly a matter of usefulness in moving onward—and thus a grade that interferes with such advancement may be interpreted as lacking worthwhileness after all. This also suggests that perhaps part of the common conception of worthwhileness is grounded in *usefulness* as opposed to *inherent value*—and it is perhaps this near-equating of usefulness with worthwhileness that guides the understanding of some teachers, administrators, and parents when they judge the worthwhileness of grades—and, indeed, of school itself.

Establishing Trust and Trustworthiness

Trust is an implicit element in grading, and its place in the theoretical framework can be seen as foundational, since both the assumption of truth and that of worthwhileness are dependent on a relationship of trust between the teacher and the students. Jackson et al

discuss trust in terms of "expressive morality," meaning that trust is displayed in classrooms through remarks, facial expressions, and gestures—subtle signs that the students interpret within the context of their own classroom.

In grading, trust is established first through explanation and discussion. All the English teachers at Mellmax High explained their grading systems very carefully at the beginning of the year. Some even distributed handouts explaining their system that could be taken home and shared with the students' parents. Most also explained that changes could be made, and that these would be announced and or negotiated as the year progressed.

The test of trustworthiness came only after grades were given. However students might question particular grades, no teachers reported that their grading system was an object of distrust. As long as they held to their announced system, or warned students of changes in the system in advance, students generally trusted their teachers to grade them fairly. Changes in the grades or grading system might at first seem cause for distrust, yet the teachers claimed that the students largely accepted them—perhaps in part because they nearly always favored them insofar as they raised their grades.

One might object that, since trust and truth are next door neighbors, winning over the trust of the students enlisted their complicity in making changes to the grades that made them untrue. Yet if the teachers have acknowledged—out of truthfulness, after all—that their grading systems are imperfect mechanisms aimed at presenting a true picture of the child as a whole, students actually trust the teacher to make wise decisions about what grades should be kept and which ones should not.

Borrowing an insight from Jackson et al's discussion of "rules and regulations" (pp. 12-14), we can transfer their remarks from the way behavior rules are enforced in classrooms to the way grading rules are applied to gradebooks.

> "As simple and direct as rules sound when they are put into words, . . . they turn out to be quite complicated when we try to understand their enactment. . . . rules seem, at first, to be inconsistently enforced. . . . What gradually becomes evident, however, is that many of these apparent inconsistencies are not actually instances of the rules being ignored. Instead, they reflect refinements of the rules that are clearly understood by almost everyone present except the observer. In other words, the rule as stated turns out to be a general maxim to which there are many exceptions" (p. 13).

What is true here about behavior rules is also true about grading rules, and the students trust their teacher to adjust the rules—for them as a class, and for them as individuals—so that the grade that goes on the report card is true.

As the year goes on, the English teachers at Mellmax High regularly explain the grades they give to their students. The use of rubrics helps support the establishment of trust, and the match between commentary and grades on essays and other assignments further develops the bond of trust. The teachers are well aware that truly capricious grading will destroy their student's trust of them, and they are diligent in maintaining that trust by carefully observing their own rules.

Cheating, of course, destroys the teacher's trust in the student who cheats. And it does so for both the student who provides the answers and the one who takes them.

Giving a zero to both students is universally perceived as fair; for the one who uses borrowed answers, the penalty is based on lack of truth; for the one who provides those answers, the penalty arises from the destruction of the bond of trust between teacher and student.

Moral and Intellectual Attentiveness in Grading Systems

The preceding examination of moral implications of grading practices in terms of truth, worthwhileness, and trust sets the stage for considering how grading displays features of Hansen's moral and intellectual attentiveness. One is tempted to separate the achievement and non-achievement factors that go into grading and say that the former involve intellectual attentiveness and the latter moral attentiveness, and while there is some truth in that, the two are often so entangled that they cannot be pulled apart.

One may accept the premise of the measurement community that achievement and non-achievement factors be strictly separated in grading and that the numerical average reflect only achievement scores. Those teachers, then, who incorporate non-achievement factors into their grading schemes violate this principle of separation. However, the teachers who directly incorporate such things as attendance and participation have sound arguments for doing so—participation is an indicator of learning and may well have an important effect on achievement scores. To help students do well on their upcoming tests, it is important to know where they seem to be before they are tested, and participation provides this information. Hansen would see this as part of the "intellectual attentiveness" an effective teacher demonstrates.

Yet if one thinks of attendance and participation as a "duty" on the part of the student, and as a display of effort—and as training for the daily appearance at a regular job some

time in the not-too-distant future, then incorporating attendance and participation into the grade is an expression of moral attentiveness as well.

Grading systems, then, clearly have elements of both moral and intellectual attentiveness, since teachers, and perhaps English teachers especially, strive to be tuned in to both what the students are learning and what they are becoming.

In this chapter, we have examined ways in which practical grading issues driven by the school district's needs and the teachers' purposes contribute to the design and implementation of the Mellmax English teachers' grading systems, and considered the moral implications of using these grading systems. In the next chapter, we direct our attention deeper into the substructure of the school system to examine the moral issues that arise when non-achievement factors in grading are incorporated into the grades teachers give. These non-achievement factors include the expectations the teachers have regarding their students and the inferences the teachers draw about effort and attitude based upon their classroom observations and the quality of their students' work.

CHAPTER FOUR: EXPECTATIONS, EFFORT, AND ATTITUDE

While the school system's grade report system shapes the ways teachers set up their systems for grading at Mellmax High in fairly obvious ways—number of failures vs. passing grades; comparison with other classes and other teachers' grades, etc.—there are other more subtle influences that affect the ways the Mellmax English teachers decide how to grade their students. These involve the perceptions the teachers form about their students' levels of effort and attitude, and color the teachers' expectations of their students' performance. But they may do more than merely influence expectations—once the teachers have arrived at expectations for certain students, these may in fact affect the way in which the teacher grades the student, even if the teacher is not fully aware of this fact. The interviews with the teachers clearly revealed that the three factors of expectations, perceived effort, and apparent attitude did indeed subtly influence their students' grades.

Even more subtly, teacher expectations may influence the nature of the interactions between the teacher and the students, creating a cycle of influence and reaction that serves to keep those perceived as the better students in their superior place, and those perceived as poorer students in their inferior place. Much of the literature on teacher expectancies suggests that this may be true. Because this was not an observational study, however, this possibility cannot be inferred from this study; neither can it be disproved by it.

Early in the school year, high school teachers form impressions of their students. In order to plan for instruction and develop tests and other assignments suitable for the

classes a teacher has this year, he needs to make an initial assessment of his students' abilities and present standing. This initial assessment forms the basis of the teacher's expectations about the future performance and behavior of his classes, both as whole groups and as individuals.

All of the teachers at Mellmax High began the year with an introductory exhortation about the upcoming year. Part of this dealt with the teacher's expectations of the students, not in terms of their ability, but in terms of those things that the teacher uses to *infer* ability: attitude, effort, motivation, and actual performance. The early emphasis on what McMillan, Myran, and Workman (2002) call "academic enabling factors" is a common feature of the opening week of high school classes everywhere; how students will apply themselves in the coming year will certainly affect the outcome.

While actual performance can be said to demonstrate ability directly—and it does to some (often great) extent—teachers frequently evaluate their students' scores on tests in light of the inferences about ability which are drawn in part from previous test performance, and in part from the inferential devices of observation, classroom interaction, and one-on-one conferences or conversations. A single test or assignment tells something about a student's ability, but it may be a non-representative piece of work. Over time, as the teacher becomes familiar with her students and their "typical" work, she develops expectations about what the student can and should be able to do on upcoming assignments. In order to examine the moral dimensions of making judgments about effort and attitude and what is right to expect from a student, research surrounding these factors will be examined to provide context.

Expectancy Research and the Self-fulfilling Prophecy

Dusek (1985) asserts that "The study of teacher expectancies was born with the publication of Rosenthal and Jacobson's (1968) initial study" (p. 1). "Pygmalion in the Classroom: Self-fulfilling Prophecies and Teacher Expectations" questioned the effect of extraneous variables on teachers' grading. Conducting an experiment in which teachers were told that certain children had done exceptionally well on an IQ test, Rosenthal and Jacobson found that the teachers' expectations influenced their grading, and that grading could be influenced by many intangible influences outside the actual student-teacher interaction. Essentially, Rosenthal and Jacobson concluded that if a teacher expected a child to do well, the child was given better grades in areas such as "cooperation," "interest," and so on. Kirschenbaum, Napier, and Simon (1971) sum up Rosenthal and Jacobson's experiment:

> "…teachers were told that ..[intelligence] tests were designed to reveal students who would probably show substantial IQ gains during the coming school year. Ten children were then selected at random from each class in the four schools and the teachers were informed that these ten children had done especially well on the test. Using these children as the experimental group…an intelligence test given at the end of the year revealed that the children in the experimental groups. . . made significant gains in IQ when compared with the children in the control groups. . . . The teachers' expectations contributed to these differences. The perceived results of the first test scores stimulated behaviors on the part

of the teachers, and eventually on the part of the 'favored' students, thus resulting in a performance discrepancy" (p. 257).

Kirschenbaum et al. then go on to express their worries about the moral implications of grading against expectations:

". . . the problem is that in another situation, the stimulus might come from word of mouth (another teacher),

a look at last year's report card, the color of a students' skin, or even the clothes he wears. It might be from language or a teacher having observed the parents of a child. Who is to say what all the variables are which are impinging on the teacher's set? And to what degree are these influencing the grading process as well?" (p. 257).

Rosenthal and Jacobson's landmark study "convinced the reading public that teachers' expectations regarding the intellectual performance of children leads to actual and significant changes in performance" (Gephart and Antonoplos, 1969, p. 579) and made the expression "self-fulfilling prophecy" standard jargon in teacher education programs everywhere. But the study quickly came under attack by other researchers (Thorndike, 1968; Barber et al., 1969; Buckley, 1968, Jensen, 1969; Elashoff & Snow, 1971)) who faulted its methods and questioned its conclusion. As a result, the study of teachers' expectations and the ways they influence outcomes—and vice versa—became a significant area of research for at least the next twenty years. Several important areas of the relationship between teacher expectations and student outcomes have been explored, and, because teachers ultimately express their understanding of a student in a grade, these

will be discussed briefly to provide background for the discussion of the moral issues tied to expectations and grades.

Learning Theories and Expectancy

The earliest studies of expectancies were conducted on animals, and were tied to behaviorist conditioning models such as B. F. Skinner's. Tolman (1932) is credited with developing the concept of expectancy and, over a span of nearly three decades, wrote on the topic as it applied to learning. While his own experiments dealt chiefly with animal learning, his definition of "learning" was transferable to humans, and his work led to similar work with human learning.

> "He believed that learning consisted of the acquisition of information (expectations) concerning the outcomes of various responses. . . . The role of outcomes (reinforcements) was to be understood in terms of confirmation or disconfirmation of expectancies, not the automatic strengthening of connections between stimuli and responses" (Zuroff & Rotter, in Dusek, 1985, p. 13).

This movement away from stimulus-response theories of behavior gave rise to Lewin's (1935, 1936, 1938) "level of aspiration" research, which in turn led to research in "goal-setting, reactions to success and failure, and persistence" (Zuroff & Rotter, in Dusek, 1985, p. 18). Social learning theories incorporated these concepts and developed rapidly, beginning with Rotter's (1954) social learning theory, and leading to Mischel's (1973) cognitive-social learning theory, which sought to synthesize social learning theories and cognitive psychology (for a detailed discussion, see Zuroff and Rotter, in Dusek, 1985, pp. 22-23).

Bandura's (1977) social learning theory presents two kinds of expectancies important for the classroom: an outcome expectancy and an efficacy expectancy. The first is defined as "a person's estimate that a given behavior will lead to certain outcomes," and the second as "the conviction that one can successfully execute the behavior required to produce the outcomes" (p.193). Thus students have both outcome expectancies and efficacy expectancies—they believe doing this or that assignment will lead to this or that grade, and they believe that they can either do this or that assignment well, or not. Likewise, teachers form both kinds of expectations, for both their classes as groups and for individual students.

Which comes first—expectations or abilities?

Hall and Merkel (in Dusek, 1985) observe that "If teacher expectancy was a widespread antecedent of individual differences in student intelligence and/or achievement, then we should find that teachers do indeed treat students for which they have differing expectancies in a different manner. In effect, the high-expectancy students should be treated in a more pedagogically sound manner. The assumption [among researchers] was that teachers would decide which students they believed had the most learning potential and then teach them in a superior manner" (p. 78).

Hall and Merkel go on to summarize several studies aimed at discovering if teachers, having formed expectations of their students, then treat them accordingly. Rist (1970) concluded that they did; his study suggested that the teacher he observed gave equal "supportive statements" to all groups, but used considerably more "control statements" when dealing with the students for whom she had lower expectations.

Brophy and Good (1969) developed an observation instrument for such studies, and its subsequent use for several studies helped make replications more comparable. In a subsequent (1970) study, Brophy and Good observed that high achievers received more praise, while lower ones received more criticism, and higher achievers' answers were rephrased for the class, whereas lower achievers' answers were not followed by such feedback.

Weinstein (1976), however, found that while teachers differentiated instruction based on expectations, no consistent bias toward one group or another seemed to be present once the school year was underway, although praise seemed more regularly directed at those for whom the expectations were lower.

Cooper and Baron (1977) hypothesized that high expectancy groups would receive more praise and criticism than low expectancy groups, but this was not so; while the high achievers received more praise, they did not receive more criticism, which seemed to be given chiefly in response to inappropriate behavior, regardless of the level of (academic) expectancy. This raises the question of how classroom behavior influences teacher's expectations, and whether it might influence grading because it colors the teacher's sense of the student's performance, quite apart from his or her marks on assignments.

As studies advanced, it was discovered that differential treatment was usually a response to the teacher's perception that such treatment was appropriate; while earlier studies had set out with the self-fulfilling prophecy in mind, expecting to discover that teacher's expectations were a negative force in the classroom, developing the strong and harming the weak, Hall and Merkel remark that "we have come full circle…from suggesting that differential expectations of students with different abilities were causing

differential treatment, which in turn resulted in insufficient progress. . . to the conclusion that these differential treatments are actually beneficial to all concerned"(in Dusek, p. 85).

One of the difficulties in determining the strength of the studies undertaken to determine if teacher expectations determine interaction or vice versa is that all the studies were small scale and unlikely to be generalizable; the most likely conclusion of all the research is that the matter varies considerably from teacher to teacher, and the influence of expectations on student-teacher interaction is widely different depending, perhaps, on both the teacher and the students studied.

Expectations and classroom control

Cooper (1979) discusses an interesting dimension of teacher expectations, arguing that "context and expectations influence teacher perceptions of control" (in Dusek, p. 146). Describing the "expectation communication model" he developed, Cooper (1979) explains

that

> "variations in student background and ability lead teachers to form differential perceptions of how likely students are to succeed. This obvious point was made so that the fact that the expectation communication process involved many feedback relations was acknowledged, even though only one causal sequence among many was discussed. Also acknowledged was West and Anderson's (1976) point that performances were as likely to affect expectations as expectations were to affect performance" (in Dusek, p. 146).

Cooper thus acknowledged that expectations are, in effect, tied up with many other factors, and may not be easily separated from those factors when teachers come to giving grades. However, his model showed that teacher control over interaction with students was influenced by the teacher's expectations of the students. Specifically, "Control was defined as having at least three subdimensions. . . : teacher control over interaction content (what the interaction is about), timing (when it occurs), and duration (how long it lasts)" (in Dusek, p. 146). Other studies by Cooper (1977), Entwistle & Webster (1972), and Sarbin & Allen (1968) supported the contention that teachers' expectations of their students influence the ways in which interaction with students is initiated and controlled; this has implications for grading in that it suggests that to some extent, teachers' expectations affect what they actually ask of students, and this in turn affects the students' performance. A demanding task required of a student expected to do well may be assessed in subtly different ways from the same task assigned to one expected to do poorly. Indeed, if similar products result, they may not be graded in the same way, despite the use of a rubric or model, which, after all, has room for subjective assignment of points within some range.

What Cooper (1977) calls "differential feedback contingencies" for high or low expectancy students results not only in different interaction between teacher and student, but also in differing student beliefs about self-efficacy.

"High expectation students may be criticized when the teacher perceives them

as not having tried and may be praised when efforts are strong. Low

expectation students, however, may be praised and criticized more often for

reasons independent of their personal efforts, namely the teachers' desire to

control interaction contexts. Greater use of feedback by teachers to control interactions may lead to less belief on the part of students that personal effort can bring about success" (p. 401).

A follow-up study by Cooper and Good (1983) found that "academic praise to low-expectation students was more strongly related to the attribution of 'following the teacher's directions and instructions' than was praise to highs" (in Dusek, p. 150).

While the English teachers at Mellmax High were not observed for interaction and control, Cooper's remarks about differential treatment based on teacher expectations of students and their abilities in terms of feedback, praise and criticism can apply to grading issues as well. Many of the teachers indicated that feedback—positive or negative—was tailored, at least in part, by the teacher's expectations of the student.

Attribution and Expectations: Effort or Ability?

Attribution theory looks for the perceived cause of an event or a performance, and seeks to explain how people come to assign the cause for a given outcome. Early theorists (Heider, 1958; Kelley, 1967, 1972; McArthur, 1972) posited that people make attributions concerning behaviors or events based on both internal and external factors. For our purposes, this means that some part of the answer to the question, "Why did this student do this assignment this way?" comes from what the teacher has observed of the student, and part comes from within the teacher—based on expectations, personal responses, and perhaps, feelings (positive or negative) toward the student.

Weiner et al. (1971) arrived at four attribution categories that could be seen as the causes of success or failure. These were ability, effort, task difficulty, and luck.

For ability, an inference is formed based upon past performance, so that steady performance at a given level results in the formation of an ability expectation. Effort is an inference about how hard one tries. Jones et al. (1968) observe that effort is sometimes based on past record of performance, with consistent success interpreted as showing high effort, and inconsistent success or failure interpreted as showing less effort. Weiner (1980) argues that external appearances of effort (exertion, struggle, etc.) also contribute to the inference of effort.

Task difficulty and attribution is again both external and internal; if an assignment results in poor performance by many students, it is interpreted as difficult for all; if most succeed, it is interpreted as an easy task. But for students who do not perform well, the attribution may be that the task is too difficult for them, or that their ability is low, or both.

Luck is the least likely attribution for success in most high school English assignments. This is because the assignments do not usually involve an obvious element of chance—unless they are multiple choice tests or quizzes. However, Borko and Shavelson (1978) found that if teachers were given information about good past performance by students who were doing well, they attributed present success to ability and effort. If the teachers were given information about poor past performance by students who were doing well, they often attributed the success to luck.

The reliance of teachers on patterns of past performance—the "consistency hypothesis" (Peterson and Barger, in Dusek, p. 166), is supported by Rejeski and McCook (1980) and Cooper and Burger (1980). This suggests that the establishment of expectations based on

patterns of success or failure can color judgments about future success or failure—and may, in turn, color the grades given for such successes or failures.

Research throughout the 1970's and 1980's (Covington, Spratt, & Omelich, 1980; Cooper & Burger, 1980; Medway, 1979, Meyer, 1979; Silverstein, 1978) established that teachers' attributions of success as a result of effort are highly predictive of teachers' feedback to students. This does not, however, suggest that teachers form expectations of success capriciously. Nevertheless, a pattern of performance may lead to expectations that influence grading in very subtle ways. Weiner and Kukla (1970) found

> "that the greater a student's success, the more positive the teacher's feedback. Students who were perceived as expending effort were rewarded more and punished less than students perceived as not trying. However, regardless of the attributed effort, low ability students were rewarded more and punished less by teachers than high ability students. Effort was a far more important determinant of reward and punishment than ability" (cited by Peterson and Barger, in Dusek, p. 175).

One of the difficulties in examining how teachers may make use of their expectations of students' performance is the overlap in real life between effort and ability; for many teachers, the two are so closely connected that distinguishing between the two is difficult and one is often equated with the other. Peterson and Barger (1985), sum up the difficulty:

> "…although attributions to student effort and ability are not orthogonal in real life, this does not prevent one from examining them and making them orthogonal in actual experimental situations. …this is the only way one can

ever separate out what effects are due to teachers' attributions to ability and what effects are due to teachers' attributions to effort. As long as the two are permitted to covary, one can never determine whether ability or effort is the more important factor in affecting teachers' behavior" (in Dusek, p. 178).

While Peterson and Barger argue that it is experimentally possible to separate attributions to effort and attributions to ability, their admission that the two are not orthogonal "in real life" highlights the complexity of determining how—or whether— these attributions play out in grading, especially since most teachers—and certainly those at Mellmax High—seldom made a clear distinction between the one and the other. Further complicating the issue is the fact that both effort and ability depend, to a great extent, on perception and interpretation, so that either or both may be inaccurately assessed in any given situation.

Encouraging Effort: Motivation

Natriello and Dornbusch (1984) hypothesized "that teacher perceptions of current student effort were likely to affect the teacher's behavior, which is often directed toward encouraging future student effort. Not only would evaluation affect effort, but effort would affect patterns of evaluation" (p. 145). Encouraging student effort is a matter of motivation, and frequently grades are seen as a device for motivating (Karmel, 1970; Glasser, 1971; Evans, 1976; Hargis, 1990; Brookhart, 2004). While the obvious effect of low grades on motivation is negative—"zeroes are motivation killers" (Dockery, 1995, p. 34)—the reverse is a common assumption; i.e., higher grades increase motivation and encourage effort. The difficulty, however, is summed up by Natriello and Dornbusch:

". . . the world of the classroom is seldom a world of perfect information. Not only are teachers often unsure or mistaken about certain student characteristics, . . . but they are also likely to be unsure about student performance or behavior. Teachers appear to base their evaluative behavior on honest attributions of student effort, but their bases for making such attributions are often unsound" (p. 145).

Thus, interpretations of student effort may be incorrect, and affect grades in ways which do not, in fact, translate into improved motivation at all. Lang (2004) muses over the difficulties in predicting how a grade may motivate or de-motivate a student, and concludes that college faculty seldom know their students well enough to predict the effect of a grade given for the purpose of improving motivation. A further problem arises when one considers Hargis' (1990) observation that ". . .a poor grade can nudge some students to better performance. However, this can only be done if the student is in fact capable of a higher level of performance" (p. 9). Poor grades given to students for whom "there is a discrepancy between…ability and the work…required…will not motivate these students; poor grades will demoralize them" (p. 9). Ironically, says Hargis, since good grades are the ones that are motivating, "the only students who are motivated by grades are students who are already getting goods grades" (p. 9). While this may oversimplify the complexity of the issue, the point is worth taking—especially since it seems that this encourages teachers to give more consideration when leeway is available in grading to the weaker students, on the grounds that they will be better motivated by a few more points.

Poor grades, according to Hargis, "indicate poor practice and inadequate comprehension" (p. 22). Poor grades are more likely to indicate these problems than "a willing poor performance on the part of the student. When this is the case, poor grades cannot be motivating; they will simply compound the problem" (p. 22). Hargis goes on to argue that poor grades contribute to learned helplessness (Grimes, 1981) and, in conjunction with the motivating effect of good grades on those already getting them, to the "Matthew effect" (Stanovich, 1986), where "those who much have will receive more, and those who have little will lose what little they have."

Schrag (2001) examines issues of justice and grading in a consideration of effort in terms of egalitarian principles, based on John Rawls' Theory of Justice (1971). He distinguishes between "luck egalitarians" and "democratic egalitarians," explaining that for luck egalitarians, "those factors that are within a person's control ought to count far more than those that are not, and they agree that talent is largely a result of what Rawls calls the 'natural lottery'" (p. 65). For democratic egalitarians, the aim of egalitarianism is "not to ensure that everyone gets what they morally deserve, but to create a community in which people stand in relations of equality to others" (p. 66). In either case, the egalitarian sees inequality—whether of talent or ability, or in terms of relations with others—as a problem, and Schrag unites this issue with the issue of awarding grades.

In doing so, Schrag strikes upon two important dilemmas of grading that are certainly applicable to high school English. First, how do teachers deal with the obvious difficulty that some students can easily write an effective essay with little effort, while others struggle heroically to produce a mediocre product, and grades that reward talent over effort (or vice versa) may seem unfair to some students, no matter how they are assigned.

Second, how can a teacher guarantee that those students whose grades are inferior to others are not also made to feel socially inferior as a result? While the teachers at Mellmax High were almost certainly unfamiliar with the philosophical musings of the egalitarians, they were clearly aware of both of these problems, and struggled to make their grades account for both effort and talent while being perceived—by both themselves and their students—as fair.

Schrag sums up the chief difficulty in managing such an undertaking succinctly: "The principal difficulty facing effort-sensitive reward policies is that it is impossible to identify the proportion of a product or performance that is the result of effort alone. All products or performances are compounds of effort and trained capacity, with capacity resulting from. . . luck and choices"(p. 70). Here, luck means both fortunate happenstance and natural talent—that is, whatever innate aptitude one may have for a task without training. How the teachers at Mellmax High deal with these issues is explored below.

Attitude

Attitude is another difficult construct in the constellation of factors influencing grading. Many of the teachers at Mellmax High mentioned "attitude" when talking about factors that influenced their grading. They were not speaking of their own attitudes toward their students, but their students' attitudes toward the subject, and assignment, the course, their classmates, or the teacher. Exactly how attitude was evaluated varied from teacher to teacher, and it was clear that "attitude" was an inexact concept having considerable overlap with effort and expectations. "Good attitude" would translate, one would suppose, into strong effort; "Poor attitude" into weak effort, and "Bad attitude"

into misbehavior and trouble-making. However, it was not clear whether perceived effort was the measure of attitude, or vice-versa, and it varied widely among the teachers at Mellmax High.

Rotter (1967) describes "attitude" as a type of expectancy, that is, a person's privately held expectancies about an object, and thus attitude can be seen as an expression of expectations on the part of the student regarding a particular class or assignment. This makes separating attitude and expectation difficult, but for purposes of discussion here, "attitude" will refer to the students' apparent expression (or lack of it) of expectations about the activities undertaken in English class, *as interpreted by the teachers observing the students.* Here the difficulty of separating the two concepts becomes even more apparent, since the teacher's expectations of the students can be partly an attitude (i.e., a privately held expectancy) and partly an overt, and explicitly announced, expectation.

A student's attitude is often inferred by her teachers based on remarks in class, level of attention or inattention, behavior, and even dress. Subtle judgments about attitude reflect personal biases as well as experience and comparison with similar students encountered in the past. Viadero (2003) examined research on avoidance strategies and self-handicapping behaviors among high school students, and concluded that classroom climate and a level of expectations based on success rather than attempts at success discouraged trying and encouraged self-defeating behaviors—which are typically interpreted as "bad attitudes."

The literature on attitude is scant, undoubtedly because the construct is an unclear one and tangled up in expectations, effort, motivation, and behavior. Yet "attitude" clearly emerged as a factor in grading among the teachers at Mellmax High, and how it played

out in making decisions about grades sheds light on the moral dimensions of grading with regards to expectations and perceived effort, and all things inferred from such sources. Schematically, the relationships among the concepts of expectations, effort and attitude can be summed up as shown in the following diagram:
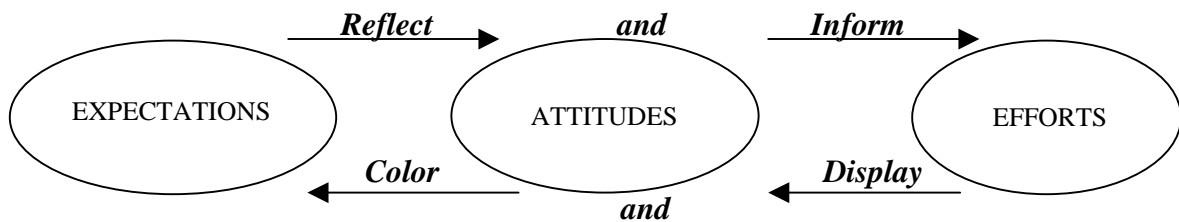
Teacher expectations reflect perceived student attitudes and both reflect and inform students' efforts; student efforts then display students' attitudes and color their teachers' expectations. In addition, expectations demand certain attitudes and efforts, while efforts shape attitudes and expectations.

EXPECTATIONS, EFFORT, AND ATTITUDE, AND GRADING AT MELLMAX HIGH

All of the teachers in the English department at Mellmax High School spoke about their expectations regarding students and their performance in class, both as groups and as individuals. All acknowledged that expectations influenced grading to some extent. This was especially true after the teachers felt that they had had enough time to get to know their students well enough to have clear expectations that they believed were accurate notions of the students' abilities and attitudes toward their work. Effort was seen as a sign of attitude and expectations were based in part on past performance, perceived attitudes, and apparent effort.

Establishing a Baseline

The initial interviews for this study were conducted at the start of the school year, and all teachers were asked whether they felt all of their students were performing at grade level. All the veteran teachers responded promptly that they were not, as did all of the

established teachers. David Dutcher, who had changed to another school district with a different profile, felt all of his students were placed correctly and belonged where they were. Ellen Enders, one of the novice teachers at Mellmax High, also felt that all of her students were at their grade level this year, qualifying her answer to limit it to this year's class and implying this was not the usual situation for her. She also acknowledged that the mainstreamed students who had IEP's were not at the grade level of their peers, and some were placed in the class for socialization purposes, not academic ones.

In fact, all the teachers qualified their answers at first with reference to those students with IEP's, since this population is very likely to consist of students who are performing at a level below the grade level to which they are assigned. When asked whether students of a lower ability were graded differently, all the teachers argued that those students with IEP's were graded according to the IEP's, and that this was almost always against a different standard, one adjusted according to the student's limitations. Jeanette Jones was quick to point out that this was required. "Well, you know, it's the law. We have to grade them the way it says in their IEP. We get a lot of memos about that kind of thing. You know, they're tested differently—extra time, separate room, they can have the questions read to them or explained, that kind of thing. But it's the law." Other teachers who remarked on the ways grading was adjusted for students with IEP's worried less about the legal issues; they felt the IEP provided a guide for them, and were not ill at ease adjusting their practice based on the information they had about the student.

Indeed, most teachers were avid information-gatherers, and used the information obtained at the start of the year to form expectations about their students' future performance. "I always give a few assignments at the beginning of the year to see how

they write and where they are," reported Luann Lender. "I don't even grade them, but I read them very carefully. It helps me know what to expect, and how to go forward with the lessons for the year." This practice was almost universal among the English teachers; all felt that it was important to establish an initial sense of abilities and expectations at the start of the year. Some, however, gave grades on these early assignments, but not always for the same reasons.

"I grade the first assignments, too, so they get an idea of where they stand and how I grade. I give them a sheet to explain how I grade, and I show them a rubric for their assignments. That way, they know what I'm looking for," remarked Gail Goodwin. This practice helped in establishing not only her expectations of the students, but also in forming their expectations of the course and its grading system. This is not to imply that Luann Lender and others who did not give grades on the early assignments failed to communicate their expectations for performance to their students. Luann returned the papers with narrative commentary and then discussed them with the class. "I make a list of all the errors I see as I read the first papers, and then I put the most common ones on the board. I tell them [the students] that they may not have made all of these mistakes, and no one made every one of them, but they should pay attention to the lesson so they don't repeat the ones they did make, and don't make the others later."

Other teachers gave grades on these early assignments for a far more pragmatic reason: the first five-week progress report required a grade. The novices, especially, felt that they needed to grade the first assignments for the purpose of having a grade for the first report. "I try to give them something I can grade, like a few reading quizzes and an essay or something because we have to have grades," remarked Alice Andrews, "But I try

to make them pretty simple so they do OK." Betty Browne hoped that early successes would improve motivation. "I try to make the first few things something they can all do well on, so they have a good start and they feel good about themselves. You know, some of these kids have never had success in English. This way, they get a good start and maybe they can do better this year."

The more seasoned teachers, however, felt that the diagnostic purpose of the early assignments was important for sizing up their students and planning for subsequent instruction; if the assignments proved to be challenging for the students, they would go ungraded to relieve the anxiety of a "bad start" to the school year. But waiting until later in the year to establish expectations was seen as wasting time. "I like to know where these kids are as soon as possible," Hester Hypoint remarked, "so I know where to go from here."

Establishing initial expectations often required the first four or five weeks of class; one of the chief objections to the first "five week progress report" was that it never actually came after five weeks, but was usually required after only three or four weeks, depending on the school calendar in a given year. "We really only had three weeks of classes, and that's not enough time to really know where the kids are yet," complained Catherine Carney, "if we use what we have right now, the grades will be different by the time they get the reports." This complaint about how the system drove early decision-making was echoed by nearly all of the other teachers, although the veterans seemed most resigned to the inflexibility of the system. Hester Hypoint tells her students, "Well, . . . these are their grades right now, but they could be totally different after the next assignment." While all of the teachers acknowledged that establishing their expectations

of future performance was very important, each went about doing so according to their own pattern, and most acknowledged that the first grades given were at best mere approximations of what they—or their students—could expect on subsequent grade reports.

Different expectations, different grading?

Once initial expectations were established, the teachers could decide how—or if—to adjust grading practices based on these expectations. The following chart indicates the answers given to the questions, "Do you grade weaker students differently?" and "Do you grade stronger students differently?"

Teacher name	Do you grade weaker students differently?	Do you grade stronger students differently?
Alice Andrews	yes	yes
Betty Browne	IEP's only	yes
Catherine Carney	yes	yes
David Dutcher	no	yes
Ellen Enders	IEP's only	yes
Fiona Fisher	IEP's only	no
Frank Fender	no	yes
Gail Goodwin	no	no
Hester Hypoint	no	no
Jeanette Jones	yes	no
Karen Kistner	yes	no
Luann Lender	no	no
Mary Minton	yes	no

The table shows that the teachers clearly adjust grading based on their expectations for performance. Earlier conversations made it clear that the teachers spend considerable time and effort to establish expectations for individual students that they believe are accurately grounded in previous performance and inferred ability. But once these expectations are established, the standard against which the student is graded reflects, at least in part, the limitations that the expectations have established.

It would seem that any teacher that answered "yes" to either of the two questions would have to answer yes to both, since grading one group of students by a different standard would mean that both groups were graded on different standards. Yet it was clear from the responses that this is not how the question was understood. The teachers understood it to mean, "Do you have a standard by which you grade, and if so, do you grade weaker students more leniently and stronger students more strictly against this standard?" Understood this way, it becomes clear that many teachers applied their standard more loosely for the weaker students, and some more strictly for the stronger students. This understanding on the part of the teachers was made clear when David Dutcher was asked about grading against his expectations for his students.

"After the first few weeks of classes, I have pretty much decided what I can expect from my students, and I have a sense of who are the better students and who are the weaker ones. Do I grade the weaker ones by a different standard? No. That would be wrong. I would be misrepresenting what they can do; I would be telling them they can do something they can't. They're in the grade they're supposed to be in, and they are supposed to be able to do the work at this level…if they can't, I have to give them lower grades.

"So do I grade the stronger students by a different standard? Well, I'm not sure what you mean. Do I grade them harder? Do I expect more from them? Well, yes, I do expect more from them. But I don't grade them by a different standard. I hold them more strictly to the same standard. I don't hold them to a stricter standard."

While it may seem like splitting hairs, what David was saying conforms to what has already been observed about the ways teachers adjust for differences in perceived abilities—the standard is largely the same, but its application is looser for the weak, and stricter for the strong students. Undoubtedly this is partly to prevent the appearance of having multiple standards, which undermines the apparent fairness of the teacher. But it is certainly true that even a strict rubric has some room for play with the assignment of points, and none of the teachers who said they graded some students differently felt they were using a double standard in their grading.

Some interesting patterns in this practice emerge if one looks carefully at the table and compares it with the earlier table that specified grade levels taught and years of experience.

First, none of the veterans admitted to holding the stronger students to a stricter standard. Luann Lender explained, "I've been teaching for many, many years, and I know what to expect. I set my standard where it should be, and then I grade my students against that. I don't need a higher standard for the better students. They know what to expect, and what they need to do to get a good grade. It wouldn't be fair to say, 'I expect something more from you because you're brighter.' If they do what they're supposed to, and they do it well, they get the best grades."

Hester Hypoint confirmed this position, saying, "I know what I can expect from a good ninth grader. If they give me what I ask for, they get the grade they deserve. I don't expect more than I ask for." Years of experience gave the veterans greater self-assurance that they had appropriate standards for the work they asked for; the novices were less sure, and perhaps sometimes felt they had not asked for enough when they reviewed submitted assignments. Or they may have discovered, based on the assignments received, that they could have reasonably expected much more than they had anticipated, and then felt the standard they had created for judging the work was too low—and hence, felt it justified to grade the stronger students differently from the bulk of the class. Of course, we have already seen that the novices often threw out assignments that proved to be of little worth, and adjusted their entire systems for grading over the course of the year, so this kind of adjustment does not seem surprising.

However, why would three of the five veterans grade students of lower ability differently from the rest of the class? The three who did, Jeanette Jones, Karen Kistner, and Mary Minton, all taught 7th and 8th grades, whereas Hester Hypoint taught only 9th and Luann Lender only 11th. The junior high years are the transition years at Mellmax High, and the students come in from the elementary schools with different ideas about schoolwork and grading. "They [the elementary students] don't do any homework," complained Jeanette, "all the work they do is done in class. What they don't finish doesn't get done. And if they have a teacher who doesn't like to read, they might not do any reading for two or three years in a row, depending on which teachers they get." This meant that for many of the students, working independently was a new practice, and being graded against set standards was, by and large, an unfamiliar idea. So the teacher

who worked with the junior high students, especially the 7th grade teachers, felt it was unfair to hold all the students to a standard that they had yet to become accustomed to. In addition, the junior high teachers invariably complained that many students came to them unprepared at a level well below what they expected, and worried that low grades against strict standards would demoralize them, causing them only to fall further behind.

Karen Kistner argued that, "My classes are always a mix of many, if not most, below-level students. Some have IEP's but most are just very far behind. There are a few students who are at the right level, and I feel sorry for them sometimes, because their classmates are so far behind, and a lot of the work we do is for them. But most of their grade comes from things we do in class, [she called this homework, but it was never taken home to be done there] and the better students can do the other assignments that we have them take home to do. They'll do better at those than the weaker ones, but the work the weaker ones do in school will keep them from failing. If it depended on the work they did all on their own, they'd all fail, except those two or three 'good students' that really don't belong here."

The only junior high teacher who did not admit to grading weaker students by a different standard was Gail Goodwin, whose approach to grading was remarkably similar to David Dutcher's in the tenth and eleventh grades. "I give them the rubric and I tell them, 'Here's what you need to do to get an A, and here's what you need to do to get a B, and so on. I don't grade anyone differently. They know what to do, and I help them learn to do it." She argued that rather than adjust her standard, she would seek out help for the weaker students, especially those who were "classified," meaning labeled as having special needs or IEP's. She worked closely with the special education staff and

shared all assignments—and their accompanying rubrics—with them so that they could help the students with whom they worked to hand in work closer to Gail's expectations.

Which come first—expectations or grades?

Clearly, one of the first challenges a teacher faces with a classroom of new students is the process of "sizing up" the new class. This is one of the first tasks of the new school year, and all the teachers at Mellmax High took it very seriously. Initially, expectations are based on what the grade level requires, and every teacher began the year by explaining what they expected *in general* from all the students. But a good deal of this had less to do with performance expectations in terms of cognitive display, and more to do with rules and regulations, when to have assignments in, how homework should be done, whether books had to be brought to class everyday, what sort of notebook to have, and so on. All of these kinds of things helped to indicate attitude and effort, and at least initially, this colored the teachers' judgment of their pupils and shaped their expectations.

"I have a lot of good kids this year," remarked Gail Goodwin, "they know what I expect and they come prepared." Were they at grade level? "No. Not all of them. They never are. But I tell them what I expect, and we go from there." Some of Gail's early assignments were graded more heavily for being in on time than for content; the purpose was to establish the importance of handing assignments in promptly. Points were lost for late assignments, and within a few days, it was not worth attempting an assignment. But as the year progressed, the rubrics used for grading, while still including penalties for lateness, shifted the weight more and more toward the performance demonstrated in the assignment. Thus, grading practice shifted subtly over the course of the year, moving

from a base of expectations grounded in the classroom rules to one increasingly grounded in academic performance.

This was especially true of the junior high teachers and the novices; for the senior high teachers and the veterans, levels of expectation with regard to performance were made clear early on and assignments were graded to determine what could be expected in terms of the already established expectations for the grade level, based on the logic of sequential development (for the non-veterans) and on both the logic of sequential development and experience (for the veterans). This often meant that the veterans had a clearer picture of what to expect; years of dealing with gaps in the sequence, so to speak, had made them better at predicting likely areas in need of improvement, and they spent little time complaining about missing steps in the development of their students.

Initially, then, expectations of one kind—general and grade-level specific—guided the teachers in both giving assignments and in grading them. As the students began to fall out into patterns of performance and the teachers could begin making comparisons between and among their students, their expectations for individual students became clearer and more personalized. After the teachers began to feel that they were familiar with their students, expectations of a second kind—particular and student-specific— guided the teachers in interpreting assignments and giving grades.

Now that I know you…

The shift from generalized expectations to student-specific expectations was made clear in the interviews conducted at the midterm, just before school let out for the winter holidays. At the start of the year, each of the teachers was asked, "Do you separate the students' identities from their papers?" In the fall, nearly all the teachers (the exceptions

were two of the longest serving veterans, Jeanette Jones and Mary Minton) replied that they did—or at least, tried to. Ellen Enders gave a typical response:

"Yeah, I try not to look at the names when I'm grading them. I don't cover them up or something, but I try not to look at them. I'm trying to judge them against the rubric. Is this a good essay or a poor one? It doesn't matter who wrote it. I want to be fair, and not let 'whose paper this is' make a difference."

Hester Hypoint put it differently. "It doesn't matter who wrote it. I'm grading against a standard, and if this paper has what it takes, it's an A, or a B or whatever. If it doesn't, then it gets an F or a D. I keep personalities out of it."

David Dutcher echoed Hester's comments. "I use my rubric and I mark everybody the same. I make it clear that this is what I expect, and this is how they'll be graded. It doesn't matter whose paper it is."

Jeanette Jones and Mary Minton, however, responded differently. Jeanette said, "You can't really separate the kids' identities from their work. How do you see if there is a pattern, if you don't know who it is? And how do see progress from one thing to another?" Mary Minton remarked, "I don't know if you *can* separate the identities from the papers. Maybe you should. Maybe if it were an ideal world, we could. But within a few weeks, you know who's who, and you're looking for certain kinds of improvement with certain kids. So you have to know who's who."

At first, the difference in answers between these two veterans and the rest of the teachers suggested very different perspectives on grading. However, by the time of the midterm interviews, when the question was posed again, many of the teachers gave

answers closer to those of Mary and Jeanette, while a few (notably Hester Hypoint and David Dutcher) held to the same answer they had given earlier.

Betty Browne, Catherine Carney, Ellen Enders, and Fiona Fisher all recalled the question from the earlier interview. "You asked that question before, didn't you?" Betty asked, "and I gave you a different answer." Indeed, she did. Whereas her earlier answer, like those of Catherine, Ellen, and Fiona, had been that she tried not to know whose paper was whose at the beginning of the year, now she said, "Of course I don't separate their identities from their work." Catching herself in the contradiction, she now offered an explanation. "Now I know who they are. So I have different things I'm looking for for different students. Some are really strong in one area, say writing, but not in another, like taking tests. Others have mistakes they're always making that need to be corrected. So I need to see if they're making progress and doing what they need to do. At the beginning of the year, I didn't know them, and they didn't know how I graded. Now we know each other, and they know I know whose papers are whose." All the teachers who gave a different answer at the midterm had a similar rationale for the change: at the start of the year, expectations were based on an ideal student for the grade level; now they were based on actual expectations for individual students.

For the holdouts, David and Hester, who continued to grade against the ideal standard, more time was spent in adding commentary to assignments to encourage students to keep trying despite poor grades. Hester allowed students to re-work essays that were deficient, and tests that depended on memorization (vocabulary and spelling) could be retaken. "You know how to spell or you don't. Once you know, prove it to me, and I'll give you the credit. But don't be taking tests over if you don't study for them!"

David did not allow re-takes; he considered it a matter of character development that students should be prepared on time. Illness and family misfortune could postpone a test, but failure to prepare could not. And essays could be submitted in draft form in advance, not again after the fact for revision: that, David felt, was unfair to other students who prepared their assignments conscientiously and submitted them as required.

Balancing Identities and Expectations for Grade-level Performance

Clearly, once individual expectations were grounded in performance patterns and the teachers felt they knew their students well enough to judge each new work against both a general standard and a personally tailored one, difficulties in dealing with both at once arose. These quandaries were resolved differently by different teachers, but not always in a fully conscious way. Frank Fender presents a case in point:

"Do I take the individual student's past performance into account when I grade their papers? With writing sometimes, I think you have to. You can't help it. Because if they are capable of doing something and they perform, um, either better or worse, you have to take that into account. It's not a measure that I think of when I do the grading, but I certainly have those expectations in the back of my head."

Asked, "What do you do with an assignment that's a disappointment to you?" Frank responded, "I'll usually let the student know. Even before I put a grade on it, I'll ask 'em, 'What happened on the assignment?' And usually I'll get an answer, 'Well, I didn't spend enough time on it.' They'll basically take the blame—that they didn't do their best. I'd say that ninety percent of the time that's the case. The other ten percent of the time it's that they just didn't understand the assignment or they're just not good at that kind of assignment, whatever the assignment is, whether it's writing or whatever.

"I try to let them know that the grades are important, that the names of the students are meaningless to me when I'm reading the work or grading the test. I have certain expectations of them, and the grades are whether they meet them or not." But asked whether he could separate the students' identities from their work, Frank replied, "I try to…I don't know…I think I do a good job of it…there are occasions where you can't separate them, and then it's part of the grade, but usually I try to."

The difficulty of explaining how the personal expectations Frank had for individual students—kept "in the back of his mind"—influenced the grading was one that all of the teachers struggled to articulate. Most treated the effect of personal expectations as rare incursions into their grading, but allowed that it happened, especially in the extreme cases where a very good student turned in inferior work, or a rather poor student handed in surprisingly good work. Yet the typical reaction in the latter case was that the student was cheating, and this called for investigation.

"You know when a weak student suddenly hands in something they have never done before that they probably had help or didn't do it themselves. So you have to ask them to come in and talk. Usually you'll find another paper just like it by a different student, and then you have to talk to both of them," said Catherine Carney, "Sometimes I think they think we're stupid." Too dramatic improvement too suddenly simply did not fit in anyone's expectations, and so led to trouble for the student who suddenly re-shaped him- or herself in terms of performance. The only time when this was possible was if the student had simply never handed anything in before; then a stunning essay from out of the blue was accepted as genuine—but the expectations that followed were high, and there was no going back. "I've had students who just decided it was time to show what

they could do," mused Luann Lender. "After two marking periods of doing nothing, they would suddenly turn in the most fabulous essays. It was great—but I always wished they would've started sooner." This was partly because the system required an average of all the marking period grades, so very low marks in the early quarters would be averaged against the "true" marks being earned in the last quarters—giving the impression that the student was, at best, only a lackluster performer.

As for the good student who handed in inferior work—whether against the general standard or against the personal standard he or she had helped to create in the mind of the teacher, such poor work was usually attributed to lack of effort, especially among the teachers at the senior high level. Junior high teachers were more willing to allow for misunderstanding of directions, inability to do the assignment due to inexperience, or interference from family situations or personal problems.

"I am so disappointed when some of the brightest students just won't put the effort into their work. I let them re-write an assignment if they are unhappy with the grade," said Luann Lender, "and sometimes they will re-do them and turn in the best essays they have ever done. But I would say out of a hundred students, only two or three will even bother to try it. As long as they passed, they don't see the point in doing it better. It's a shame."

Mary Minton admitted quite frankly that she graded students of lower ability by a different standard. "They just don't know how to do some of the things we ask them to do. They didn't always get the background they needed in grade school, and it wouldn't be fair to ask them all to perform at the same level right from the start. So you grade the

weaker ones a little more leniently until they get better at it and come up to the level where they need to be."

Jeanette Jones argued that many of the weaker students had home lives that militated against success, and she felt compelled to take such things into account. "You know, we take a lot for granted about these kids. But some of them have never seen their parents read a book. They don't have any books at home. Nobody reads. Then they come to school, and we tell them, 'Take this book home and read it.' They don't know where to begin. You know, they really aren't where the other kids are. So you can't grade them the same. They need to be encouraged. If all they can get are zeroes, they'll just give up." Her tactic for addressing this issue, however, was not actually to adjust the student's grades, but to offer second chances or substitute assignments for missing ones. Recall that her system of grading treated all things equally, so several minor homework assignments could easily offset a zero on a test covering a book the student "could not read."

Karen Kistner's long experience working with remedial students seemed to have given her generally lower expectations for most of her students. As a result, she focused on assignments that could be monitored in class. This helped her develop expectations that were supported by regular observation both of the students' finished work and of their work habits. Regular application in class could translate into points on assignments, and thus effort in keeping with expectations would improve her students' grades.

"I give very little outside work, that is, independent work. Most of our work is done right in class. That way I can help them and if they are having trouble, we can work it out right here. So many of them are below level that we need to work on things right

here. If you ask them to do it on their own, you get nothing." Having expectations of personal performance that had come to form the norm for general expectations, Karen adjusted her assignments to meet these expectations, even for those students who she knew could perform better.

Effort vs. Ability—Attribution and the Teachers of Mellmax High

When asked if grading were affected by effort, the teachers at Mellmax High offered several variations on the ways their grading was influenced by what they perceived as effort on the part of their pupils. Most insisted that what was graded was only the product of effort, and not the effort itself. A good paper usually resulted from appropriate effort, and a poor paper from lack of effort. This typical attribution of success to effort—and of failure to lack of effort—echoed the literature on expectations and attribution, but the actual strategies for translating effort and its correspondence with expectations into grades varied widely among the teachers. At one end of the spectrum, Hester Hypoint and David Dutcher acknowledged only that effort affected the product they received, and the assumption was that good products had high effort—or high ability—behind them. Poor papers may have taken great effort on the part of some students, but this was not to be factored into the grade because this would distort the truth: "I can't give a high grade to a poor paper simply because someone tried very hard," asserted David, "a poor paper has to get a poor grade, no matter how hard the student tried on it." Hester concurred. "These are ninth graders, and they have to do ninth grade work. If they want to get more points for more effort, they need to talk to me about doing the assignment again. Then I can show them what's missing and what they can do to improve their work. But I need to see that the work has improved." Still, Hester hedged a little when asked if the re-

submitted paper showed little real change, but clearly was a serious attempt at improvement. "Well, I may give them a few points because they really worked at it. It won't change a D to an A, but it might make a D- into a C. Otherwise, they'll never try to do another assignment over again." For David, the assignment was finished the first time; to improve one's average, one had to do better on the next assignment.

Alice Andrews had perhaps the most curious measure of perceived effort, although on the surface it was similar to David Dutcher's. She, too, expected papers to be handed in on time. Points were deducted if papers were late. However, if a low-ability student submitted a paper on time, Alice interpreted this as a special effort, and awarded points for this alone. "I try to reward effort. If a kid hands in a paper on time, and they could have handed it in late like the rest of the class, I can see they are really trying. So I give them a few more points for effort. Especially a student who is struggling, so writing is hard for them in the first place. Maybe they don't have all their spelling right, but I won't hold that against them as much if they put in the extra effort to hand it in on time." This practice was not one Alice announced to her classes; the announced policy dealt only with deductions for late papers. Her rationale was that the students should be handing in their papers on time, and if the extra points for timeliness were announced, students would be handing in things on time out of the wrong motives—not to be on time, but to get the extra points.

While none of the teachers at Mellmax High admitted to giving points for effort in and of itself, many of their homework practices amounted to just that: Frank Fender, for example, collected homework and checked it off as received. It was not graded, but the assignments counted because the average for the marking period included a percentage

for homework. This was calculated by adding up the number of checked off assignments and dividing by the total number of required homework pieces. This fraction was then used to calculate the portion of the average due to homework—essentially, these points were awarded simply for effort, since the quality of the work was never assessed, not even in terms of completion. Mary Minton and Catherine Carney, however, checked for completeness, but not correctness, thus slightly changing the measure: one could assume that a complete assignment demonstrated greater effort than and incomplete one.

Another of the novice teachers, Betty Browne, seemed to equate effort with struggle. Hard work from strong students did not seem to be the same effort as struggle on the part of weaker students. "If I can see a kid is really struggling, that it's really hard for him to do this kind of assignment, and so I can't expect too much, I try to encourage them. So like with the rubrics, I can give a 'high five' or a 'low five,' not just a five. That gives me some room to help the ones who, you know, really try, but just aren't as good at it as some of the others."

While some of the novices experimented with their grading systems to motivate the lower ability students, or at least to prevent them from becoming too discouraged, other novice teachers saw trying and learning to try, and failing and trying again, as an important "life lesson."

"When a kid tried really hard and doesn't get a very good grade, I try to talk to him and let him know that he'll be all right. You can't just give up. Sometimes you need to try again, and so I encourage them to do that. Part of succeeding in life is understanding that trying again until you can do it is important. I try to put positive comments on their papers, even if they get low grades," remarked Ellen Enders.

Both Catherine Carney and Frank Fender (one of the established teachers) used the phrase "life lesson" when talking about effort and their expectations of their students. Yet in both cases, the discussion seemed to apply more directly to the lower ability students. More was expected of the higher ability students by both of these teachers, but when a higher ability student did not meet expectations they received a comment to that effect, but were not pressed into trying again. The opportunity was offered, but whereas the lower ability student would be required to come in to discuss the weak paper, the higher ability student had the option of letting it stand if he or she were content with the outcome. Both Catherine and Frank expressed some regret that the higher-expectation students seldom bothered to do work over to improve themselves or their grades, but neither felt it was the best use of their energy to pursue them.

The junior high teachers, again, were different from the senior high teachers when it came to attribution of success or failure. Among the junior high teachers, effort was considered an important ingredient for success, but failure was more often attributed to a lack of knowledge or practice than to a lack of effort. This was true of both the veteran junior high teachers—Jeanette Jones, Karen Kistner, and Mary Minton—and the less experienced junior high teachers—Alice Andrews, Betty Browne, and Gail Goodwin. "They just don't know how to do so many things when they come to us," lamented Jeanette Jones, "you almost don't know where to begin. They're so weak. It's almost as if they haven't gone to school."

Karen Kistner seconded Jeanette's contention that junior high students were largely unprepared, and expectations had to be lower than one might have liked. "They don't really understand how to write an essay, what a topic sentence is, how to provide support

for your argument, how to organize their writing. So I try to keep it simple and help them learn. If they try, and they do what I ask them to do as well as they can, they're not going to fail."

Asked whether effort in itself affected grades, Gail Goodwin revealed an interesting difference in the way exceptional work was graded compared to work that met her standards as stated in the rubric, but was not outstanding. "the only thing I have done on certain assignments, and it's more with the quality of work, and it's not just for the good kids, it's really for anybody who applies himself, I'd give them bonus points. So you would earn *more* than—like if the assignment was thirty points, but your quality of work was *so* good, and it was *so* well done, I felt like, 'that's just not enough,' and in some cases I gave bonus points. But my kid who consistently does no homework, and the kid that consistently does homework every day, if they have the same assignment done, they both get the same points." Whether the exceptional assignment was attributed to ability or effort was not relevant; the outstanding assignment, regardless of its genesis, was worth more for its quality. In one way, this is consistent with Gail's view that a standard has to be met; in another way, however, it suggests that the standard is not actually set to capture the best work, but what can be typically expected as good work.

When discussing the consequences of having students of lower ability mixed in with those of average or above-average ability, Mary Minton expressed her feelings that keeping a standard that could be applied without adjustment was quite difficult. "It's a challenge, because you want to give the same fairness; they're in that same classroom, they need to take the same test as everyone else. You can't just give them points for doing nothing because that detracts from the top student who really is achieving. So it's

difficult, but I think that if you can give students like that—if you have enough grades so that there are things they can really study for, things they don't have to really think about real carefully, as far as like writing assignments and that kind of thing—although they got a lot of assistance with writing assignments too, and I always gave them opportunities to re-write, and to conference with me, so I think if you do that kind of thing, you can get them through."

When asked if students of lower ability were graded by a different standard, Mary paused, then answered, "I think perhaps in their writing assignments…I would always look for improvement. I always looked for improvement from all my students, and I always considered that to be part of the grade, but I think it's almost impossible not to look at those students and perhaps subjectively give them perhaps some kind of credit for effort and for—if they re-wrote a paper—and for giving you everything that they had. Because it is difficult to say, 'You're giving it everything you have, but that's just not good enough.' Certainly if they tried to incorporate the suggestions that I gave and they worked really hard, I would say…probably a slightly different standard. It would be difficult for me not to. Whether that be right or wrong, it would be difficult. . . . Did I grade the more able student harder than I did the less able? No. But I think when I looked at the less able students, perhaps I looked at their effort and their improvement more than when I looked at the more able students."

When Mary graded writing assignments, she awarded grades in two categories: one for content, and one for mechanics. She explained that the content category gave more leeway for awarding points, and while the more able students invariably did better with

the mechanics part than the less able, she found room in the content grade to "give something" to the less able, so that "they were able to do all right."

Mary also raised an important issue in grading expectations: those of the parents. Observing that as she neared retirement, more and more of her classes included children with IEP's and support systems to enable their participation in the regular classroom, she remarked that these students received considerable assistance from teacher's aides, special education staff, and their regular classroom teachers. Even with all this support, she remarked, "they were probably never going to get high, high marks," but their parents expected that they would get grades like their classmates because of the extra help and assistance. "I was surprised at the number of calls I would get from the parents of special ed kids, who really objected to their students getting low grades, because of all the assistance they were getting. But I don't think they could ever get high, high grades."

The underlying expectations that the teachers at Mellmax High formed about their students were shaped partly on the basis of interpersonal experiences with individual students, partly on general experiences during each school year, and, as we see in the case of Mary Minton, partly on long term experience teaching the same grade level for many years.

Fiona Fisher's way of presenting effort and its relationship to grades began with her first class, when she explained how students would be graded. "I don't tell them that I assign grades. It tell them that the grade that they receive is the grade that they have *earned*, because I personally don't feel that I am giving them a grade—it's based on what *they* do and how hard *they* want to work or study or do their homework, so they're *earning* the points. And also I tell them that their attitude counts a great deal, and that's

so there won't be behavior problems." How exactly attitude was calculated was unclear, yet the connection between effort and attitude in the presentation of her grading system implied that she inferred a positive attitude if a student worked hard and seemed to be putting in effort. A negative attitude was inferred if a student did little work or seemed to put forth little effort.

Thus it would seem that Fiona, like most of the Mellmax High teachers, inferred attitude from effort, and believed that better grades were the result of effort, whether this effort was actually observed or not. A poor paper reflected poor effort and therefore, bad attitude. The exception to this understanding occurred among the junior high teachers, who more often attributed poor work to low ability and depended more heavily on classroom observations during assignments done in class time to judge both effort and attitude.

Despite this tendency to infer attitudes and levels of effort from the quality of assignments, most of the teachers worried about misjudging the potential of their students. Fiona Fisher is a case in point. Remarking on the comments used on the grade report sheet, she described a dilemma mentioned by several of her colleagues.

"I would like to put the comment such as 'student is not working to potential.' Because I have seen visually…that I personally think that they can be doing better. And it's touchy because if you put down that 'the student is capable of doing better,' you always have the fear of the parent saying, 'Well, this is as good as it gets.' And perhaps they will be very offended by that. But I personally think that maybe they *are* capable of doing better, and that's why I put the comment there. Maybe I see them in class not paying attention, wasting time, or not completing the easy vocabulary assignment

because they're wasting their time. So in my mind they are capable of doing better. And again, this goes back to, because I've had kids last year and then I have them this year, I know what they're capable of doing and I know how they did last year, so I can see a big difference if they're not doing the same academically."

While Fiona could reference last year's performance when judging potential in some of her students, Ellen Enders and Catherine Carney hesitated to make firm judgments about the potential of their students because, as Ellen put it, "I only see them in English class for one period a day. They may actually be good readers or writers, but I don't see it here.   Sometimes you can't really say what potential they have. They may just not be showing it to me." Just the same, if a paper were a disappointment to any of the teachers, none hesitated to write a comment to the effect of, "I think you can do better" atop the page near the disappointing grade. If a teacher felt confident that she knew the potential of her students, she might write the unqualified, "You can do better" instead—and both Luann Lender and Hester Hypoint were confident by the midterm that they had accurately assessed their students' potential. "Oh, sure, I tell them they're not working to their potential, and I tell their parents, too. I've seen enough of their work to know what I can expect of them," asserted Hester. "You can tell who can do better by the time you've read as many papers as I have," declared Luann, "I used to spend hours every weekend reading their papers. I got to know my students very well."

While an examination of grading systems provided insight into concrete ways in which teachers allotted credit and presented explicit policies for awarding points or assessing penalties on assignments, the moral issues involved in expectations and the implicit decisions teachers make about grades as a reflection of effort and attitude are

159

more subtle. Reflecting on the moral aspects of these decisions will provide a deeper appreciation of the complexity of grading in high school English.

Applying the Moral Framework to Effort, Expectations, and Attitudes

The moral issues rooted in assessing non-achievement factors fairly are so deeply embedded that they are seldom examined: by the time students reach the high school, both they and their parents have come to take it for granted that their teachers are not only deeply perceptive, but highly accurate in assessing their students' innate abilities, true attitudes, and levels of motivation. It is only when an action on the part of the teacher seems patently unfair that his ability to accurately judge the "academic enabling factors" influencing his students' grades is called into question. This means, then, that most teachers, most of the time, are perceived by their students to be reasonably close to the mark in their judgments of effort, expectations, and attitudes. An examination of these areas in terms of the theoretical framework established earlier will help shed light on the moral complexity of making grading decisions based on these inferential sources.

Truth in Expectations

In the preceding chapter, we discussed the dual perspectives that guide most grading: that of "truth finding," and that of "truth telling." While different teachers establish their grading systems to favor one or the other, all teachers must at some points find the truth, and at other points tell the truth. Grades are the means for both of these processes. This becomes even more apparent when considering expectations in the moral light of truth—how does a teacher fairly establish expectations that meet the requirements for being true? And, having found the level of expectation that is true, how does the teacher use

grades to effectively tell the truth about the extent to which a particular student is meeting those expectations?

At the start of the school year, teachers' expectations are general and based largely on previous experience—or, in the case of novices, on their beliefs about what *should* be reasonable in terms of expectations given the planned curriculum and the overall sequence of schooling. But true expectations reflect both the requirements of the curriculum and the requirements of the students as well as those of the teacher, and teachers' earliest announcements of expectations may be very near or very far from the mark. And so teachers begin a quest to ascertain the truth about their expectations, and determine the match between what is reasonable for this year's classes both as groups and as individuals.

This process involves giving assignments, giving grades, giving feedback, and processing all of this both at the individual level and at the aggregate level. Hardly an easy task, and one that, for the English teachers at Mellmax High, required several weeks. When the first progress report was required, a nearly universal complaint was, "We don't know these kids yet."

While this was partly the result of not yet having had enough time together with the students, it was also partly a function of the large number of students with whom the teachers had to become familiar, and partly a function of the enormous amount of detail the teachers were sifting through in order to establish what they believed were true expectations for their students. In addition to pen-and-paper assignments, teachers weighed in observations made during class time and in the halls; anecdotal information from parents, other teachers, counselors, and coaches; official records from the central

office and support services; bits and pieces of personality revealed in interpersonal exchanges—all subject to interpretation, and all subject to change. And yet, every teacher at Mellmax High could confidently claim by the midterm interviews that they now knew their students and what they were capable of doing.

Having now found the truth and set their expectations, the teachers used their grades to tell the truth. A poor grade from a low ability student was an expected match; likewise, a good grade from a higher-ability student was also a match. When these did not match, an explanation was required; for the low ability student who achieved a higher-than-expected grade, there was either of two alternatives. Perhaps this is not his work; perhaps she got help; perhaps he cheated, especially if the grade was well above expectations. This first alternative would require an investigation. The other alternative, that the student had progressed, was certainly preferable, and would require an improved level of expectation, and an adjustment on the part of the teacher. To fail to adjust expectations would be immoral; it would prevent accurate appraisal of future efforts, and could negatively affect future interaction with the student.

Truth and Consequences: the Threat to Trust

For the high-ability student who handed in an inferior assignment, truth required an explanation: insufficient effort, lack of interest, haste, carelessness, or perhaps a personal or family crisis. Here, one might notice that in some ways the drive for truth threatens the bond of trust between teacher and student.

In both cases mentioned above, the initial explanations for performance-different-from-expectations yield several negative possibilities: cheating, outside assistance, laziness, carelessness. A too hasty conclusion, if erroneous, leads to a destruction of trust

between the teacher and the student. Even when the teacher is honestly attempting only to ascertain the true cause for the difference between expected work and what is received, and may conclude that the student is brighter than at first thought, or needs help with a family situation, an insensitive approach to the matter can have disastrous results: a loss of trust leads to animosity, refusal to work, and resistance to efforts at reconciliation. Every teacher dreads the possibility of erring in this way, yet all teachers find themselves from time to time in situations where the truth requires that they risk the student's trust in order to establish fair and accurate expectations.

Sometimes when a teacher cannot reconcile an assignment with her expectations, she will allow the student to do the assignment over again. This is especially true when the assignment is a disappointment in terms of quality. But now and then, an assignment that exceeds expectations is a source of suspicion, as discussed above, and, in order to maintain both the truth and trust, the teacher allows a student under suspicion to attempt the assignment again without actually confronting the student about the belief that the work is fraudulent. This allows a kind of face-saving for the student, and gives the teacher a chance to re-confirm or re-evaluate her expectations. Such situations test the expressive morality of the teacher, who may reveal by his expression a distrust of the student, a sense of disappointment, or a hardening of attitude toward the pupil, which can affect their future relations for better or worse. Some of the teachers at Mellmax High, like Ellen Enders, tried to find subtle ways to hint at suspicions without destroying trust—writing a comment at the top of the page like, "How much of this is yours?" Others, like Gail Goodwin—from whom a paper that didn't seem to match with the

expectations might get the more blunt comment, "Do your own work"—were more direct and matter-of-fact.

To some extent, keeping trust and truth together when dealing with expectations and effort was a balance that depended on the teacher's personal style and the level of trust already established between the teacher and the students. Frank Fender's remark that his interpretation of inferior work being due to lack of effort was right "ninety percent of the time" showed that he had indeed come to know his students well. That they were willing to admit that they had not tried hard enough showed that they trusted him, and knew what his expectations of them were. However, Frank was aware that he might be wrong, and provided the students with a chance to explain before he drew his conclusions. This sense of caution was echoed by Catherine Carney when she related the tale of identical papers submitted in the same class by a top-achiever and her low-ability best friend. "You have to be so careful. You really have to be sure before you say someone cheated. Otherwise, you can expect trouble." Trouble could come from either student—or both— or from their parents or the office. The more experienced a teacher was, the more cautiously he approached the issue of dealing with mismatches between his expectations and what he received from his students.

Matching Measures with Expectations

Another moral issue is raised when methods of determining teacher expectations are faulty. Some of the inexperienced teachers at Mellmax High spent a good part of the first marking period assigning things that were chiefly intended to motivate their students, to encourage effort at the outset of the year. In itself, this is an admirable goal. Yet if these assignments were too easy, they could establish false expectations for both the teachers

and their students. Alice Andrews and Betty Browne both argued that in the first marking period, assignments were things "everybody could do." In part, this was done not only to motivate, but to be sure the first period grades were good for the district report card. Neither Alice nor Betty was tenured yet, so one might detect another motive in their desire to give their pupils a "running start" on the year's average. But even if this motive were not there, some question is raised about the truthfulness of grades that reflect little need for effort, and which may actually backfire as motivators because they establish low expectations for future demands from the teacher in the minds of the students—and may, in turn, undermine the students' trust in their teachers.

Thus, the measures devised to determine expectations need to be truthful in the sense that they do not establish a standard that is lower than what is actually needed to succeed in the year—and to be prepared appropriately for the next year's work.

Worthwhileness and Effort

The need for truth in establishing the benchmarks which give rise to reasonable expectations—both for the teachers and the students—leads to a reflection on the worthwhileness of the work assigned. As part of the substructure of the curriculum, worthwhileness is assumed—why would we send our children to school to waste their time on assignments that were too easy and did not develop their skills? Effortless assignments do little for cognitive development, although they may keep lower skills like memorization sharp. True, part of evaluating actual ability levels will require giving assignments of differing difficulty, and some students will find even relatively easy assignments challenging enough that these can be used to establish future expectations for these students' efforts.   For students for whom an assignment is no challenge,

however, expectations may be set too low, and the student may progress little over the course of a school year because the effort asked of him is no effort at all—making the school year devoid of worthwhileness.

On the other side of the coin, a student who finds the curriculum too challenging faces another difficulty: constant effort and struggle yield little return, and end in frustration and failure. Here it seems that the work itself demands more effort than is worthwhile, because it ends without progress or success. Much of the literature on the cycle of failure among students is based on this problem: worthwhileness in curriculum is not only a question of the value of a given skill or piece of knowledge in itself, but also of the worth of that skill or knowledge to the individual. The teachers at Mellmax High generally believed in the worthwhileness of *most* of their assignments for *most of their students.* Yet there were some, like Karen Kistner, who questioned the worthwhileness of some standard elements of the English curriculum for some of her students. "Should all my students be struggling to read *Romeo and Juliet*, or *Julius Caesar*?" she asked. "They would be better off if they could learn to read a newspaper and understand how headlines and news summaries work."

At the same time, Alice Andrews and Betty Browne were taking this very approach, giving "easy" first assignments in the beginning of the year in order to encourage their (weaker) students. Certainly they think such assignments are worthwhile from their point of view—to motivate students—yet, if students don't think they are worthwhile, of what value are they? But, on the other hand, they may, in fact, motivate some students. This example points to the moral ambiguity teachers face: not only is there no easy way to

resolve such issues; there is no ONE way to resolve them – to be fair and truthful to some students may mean the risk of violating those same principles with other students.

Egalitarianism: Making Effort Worthwhile for Other Reasons

The "luck egalitarians" (Rawls, 1971) view ability as a gift of good fortune, an outcome from the "lottery of Nature." Therefore, it is unfair to reward those for whom a task is effortless—or, at least, it is unfair not to reward those for whom the task is accomplished only with great effort. Despite a kind of logic in this thinking, the luck egalitarians have yet to find a way to reward effort in a way that does not skew reality and make judging ability nearly impossible. But the teachers at Mellmax High had some sense that there was an element of unfairness in asking students to do things that, for some, were easily done during lunch period before class, while for others were done only after days of struggling and re-working. This dilemma resulted in a series of second chances, extended deadlines, and alternate assignments or grading approaches, not always clearly explained or evenly applied. Yet it was the attempt to make effort worthwhile, even when it did not yield the same results for all students, that compelled the teachers to devise such schemes. Sometimes, these schemes for alternative approaches to assignments and grading were unique and set up for only one student; for those students with IEP's, in fact, this was often the case, and the approach was not only uniquely tailored for the one student, but was mandated by a Committee on Special Education that worked out the plan in cooperation with the teacher and his or her support personnel.

The "democratic egalitarians" (Schrag, 2001) occupy a slightly different place in the debate over effort. They argue that grading on merit or ability is fine, since it tells who

can do what well. Their objection is to grading systems that reward those who achieve effortlessly and make those who struggle to reach second best feel inferior. This is undemocratic; everyone has their gifts to give and their part to play in the wider realities of life. And so effort should be rewarded somehow, perhaps by other means than assignment of a grade which carries messages of superiority and inferiority. Here again, without ascribing in any conscious way to a position in the philosophical debate, the teachers at Mellmax High went to some pains to offset the interpretation of a lower grade as a mark of inferiority.

"I make sure I put comments on the papers when I give them back, especially for the ones with lower grades. I look for improvement, and I try to say something positive about every paper," said Ellen Enders.

Some of the teachers saw effort and trying again as part of growing up and taking on the role of adult struggle in the real world—the assignment itself was not as important as learning to try again. "I tell my kids not to be discouraged if they don't do as well as they are used to. That's how we grow and improve," explained Gail Goodwin, "No one is perfect; we have to keep trying."

In addition to putting a different spin on the worthwhileness of an assignment by making it a test of character, many of the teachers—especially those in the junior high—encouraged their students to do all their assignments so that the things that they were stronger in could offset those they were weaker in. Hester Hypoint encouraged her weaker students to do vocabulary tests over to offset reading tests, and Mary Minton looked for assignments that emphasized basic skills and could be used to offset essays and other assignments that were often too challenging for some of her lower-ability

students. In a way, this reflects a democratic egalitarianism—everyone has her own strengths, and these should be credited wherever possible.

Points awarded for simply handing in homework can be seen as points given merely for effort; while none of the teachers who used this practice understood it in this light, the device was clearly one way of offsetting lower grades that were the result of inability so that the weaker students' averages would not be dismally lower than those of the better students. This helped prevent feelings of inferiority that would have discouraged the weaker students too much, according to the teachers who employed such policies; at the same time, the weighting of such points was limited so that it did not become unfair in the minds of the teachers or the stronger students (who, after all, also got the points for completed homework, regardless of quality).

Biases and Interpersonal Expectations

One important moral consideration in the relationship between grading and expectations arises from the very fact that high school teachers become so familiar with their students over the course of the school year. As Ellen Enders remarked at the midterm interview, "My kids know me better now, and I know them. That means a lot, because now they know what to expect and they know that I know. There are no behavior issues now, either. We've worked all that out." The development of personal relationships raises issues of bias or favoritism, especially since these can be established so gradually that neither party, the student nor the teacher, is fully cognizant of the way in which it affects both the application of the rules and the interpretation of assignments.

"Sometimes you know a kid so well that you know that they just didn't express themselves clearly. You can tell they know the answer, but they didn't put it down on

paper in the best way. So maybe you don't give them all the points, maybe four out of five, when with another student it would only be three out of five," explained Betty Browne. While this may indeed be so, it seems that there is a real danger that expectations have been colored by personal relationships, and this poses a danger to fairness, both to the other students, and to the student whose efforts are being credited as higher than they appear.

Another serious issue has already been apparent in remarks by those teachers working with the lower ability students and those with IEP's. While the teachers at Mellmax High did not hold their expectations of the students so determinedly as to refuse to take new information into account, they generally spoke of those students with IEP's in a way that bordered on the "self-fulfilling prophecy," i.e., it seemed that many of the teachers took it for granted that students with IEP's were unlikely to progress significantly in a year, and so they were graded against a different standard, not just because they had an IEP which often specified restrictions on grading and what could be counted and what could not, but because they were presumed to be unable to advance to the point of requiring that their IEP be revisited and revised.

This may indeed be a moral failing, but the way in which grading students with IEP's actually played out in the classroom may have been quite different from the way it came across in the interviews. When speaking in general terms, teachers often represent their own policies in ways that do not fully match the specific applications. One of the major difficulties in untangling the moral issues involved in grading comes from the fact that every teacher does, in fact, assess each student as an individual based on personal knowing, and this means that stated policies have many variations in actual application.

IEP's are by definition individualized, and so making general statements about how teachers deal with them risks misrepresenting the process, even if the teachers are themselves making the generalizations.

Intellectual and Moral Attentiveness

Hansen's (1999) ideas of intellectual and moral attentiveness become more evident now as we consider the ways expectations, effort, and attitude touch upon grading. Whereas a consideration of grading systems made those aspects of intellectual attentiveness that are tied to the sequence of schooling evident, the deeper, personal attentiveness that makes teaching what it is becomes more evident now.

If teachers are to accurately form expectations for individual students, they must be acutely attentive to their students' intellectual development. English teachers are responsible for developing an important skill set in their students: an ability to communicate effectively and to understand others' efforts to communicate. This involves both concrete information and abstract ideas, and requires an ability to understand both the literal and the figurative, the expressed and the implied, the overtly stated and the merely intuitive. There is no set formula for developing these skills and abilities, and there is always an element of the intangible in effective communication. In order to help a student advance in the field of expression, a teacher must know the student well, and this requires both intellectual and moral attentiveness.

The intellectual attentiveness is more readily apparent to the outside viewer—clearly, as a student progresses, the teacher must pay attention to areas of strength and weakness, and accumulate a vast array of indicators that help the teacher make decisions about the kinds of assignments to give, and the ways to interpret the results, both in terms of quality

and effort. Expectations must be established, and then adjusted as the student meets or fails to meet the expectations. Intellectual attentiveness pertains to how the student is advancing in understanding and skill, and this requires that teachers be constantly judging, refining expectations, and comparing performance against expectations. The teachers at Mellmax High School revealed how they formed their initial expectations, and how they adjusted these as the year progressed—both for their classes as whole groups, and for the individual students in their care.

Moral attentiveness begins to emerge more clearly in the grading issues of high school English now that we have delved more deeply into the non-achievement factors involved. This kind of attentiveness is concerned with how the students are developing not as *students,* but as *people*, and was not evident in the direct exposition of the grading systems themselves because they were ultimately aimed at presenting a picture of performance that is generally understood as purely academic performance. Now that we have seen how non-achievement factors affect performance, we realize that these factors, too, are subtly woven into the grades. If achievement reflects effort, then the grade reflects both achievement and the effort behind it; the teachers at Mellmax High School made it clear that if achievement is to be encouraged, effort has to be encouraged. And so they must be attentive not merely to intellectual advancement, but to moral development. They must discourage laziness and encourage application; they must foster positive attitudes and prevent students from becoming discouraged or bitter; they must praise such progress as they can find, and keep the higher ability students from becoming complacent or self-satisfied.

When the Mellmax English teachers confidently reported in the midterm interviews that now they knew their students, they were asserting an understanding that went beyond mere ability to judge an essay against a rubric. This level of intellectual attentiveness was enough at the start of the school year, but by midterms, the moral attentiveness of the teachers to their students had clearly emerged. Now they knew their pupils as people, as individuals, and could form particularized expectations that meant each pupil was being judged for effort, attitude, and accomplishment against a standard unique to him- or herself—similar, yes, to that applied to others, but tailored in keeping with the intellectual and moral attention being directed toward them by their English teachers.

In this chapter, we have seen how the moral imbues judgments about non-achievement factors which play out in subtle ways in grade reports and commentaries. The ways in which character is shaped in English class are very subtle, but every teacher and parent is aware that moral issues pervade the content of literature. In the next chapter, the embedded moral framework of schooling will be examined for its most subtle expression, the molding of character. How teachers deal with character development and achievement, and the extent to which they are aware that this influences grading—or whether it does at all—forms the content of the next chapter.

CHAPTER FIVE: CHARACTER DEVELOPMENT AND GRADING

We have looked at ways in which non-achievement factors may influence grading and how the teachers at Mellmax High incorporate non-achievement features into their grades, often without being fully aware that they are doing so. In doing so, we have seen that the substructural elements of the moral framework of schools are indeed deeply embedded, almost to the point of invisibility. Because high school English is a composite of skills, concrete knowledge, and abstract thinking, it can provide a fertile forum for the discussion of moral issues, and so it has traditionally been seen as a place where character issues could be explored. It is logical to expect, then, that English teachers' assessment of their students may be affected by the ways in which the character development of their students seems to be unfolding. This chapter explores this possibility.

From the beginning of American schools in Puritan times, the development of character in the young has been a central concern of teachers, parents, and society as a whole (McClellan, 1999). Yet, because of the unique character of American public schools and the constitutional issues of freedom of religion and the separation of church and state, school teachers have always had to approach the issue of the moral development of their students in a way that was both compatible with the general views of the local community and sensitive to widely differing ideas about what constitutes proper instruction in the development of character. This is a difficult task, and one which is made even more difficult by the fact that high school grading is, at least to all appearances, based on academic achievement and mastery of course content alone, with no apparent room for the expression of approval or disapproval of a child's character as it develops under the care of the classroom teacher.

The previous chapter, however, has demonstrated that non-achievement factors can, and do, to a greater or lesser extent (depending on the teacher—and the individual student) have an affect on grades and so do, in fact, contribute to higher or lower grades. The central question of this chapter, then, is this: can perceived growth in character also subtly influence the grading practices of teachers, and if so how? Before investigating this question, however, some context for character development and a brief review of pertinent literature will provide an exposition of common terminology and issues surrounding the matter.

CONTEXT AND LITERATURE

The literature surrounding character education is vast; character education has been a constant topic of interest in American public schools from their earliest beginnings, and theorists have promoted various approaches and methods for developing good citizens in every decade from the 1830's to the present (see Howard, Berkowitz, and Schaeffer, 2004, for a concise and fairly comprehensive overview). For our purposes, a summary of the major schools of thought and an exposition of terms will provide sufficient context to facilitate our examination of the grading practices of the English teachers at Mellmax High School with regard to the ways their efforts to develop good character in their students may influence their grading practices.

At different times, education theorists have used different terms to describe the difficult process of building character in children. "Moral education," "values education," and "character education," all share the same objective: to train children and youth in ethical principles of right and wrong, and to encourage conduct in keeping with those principles. This sounds easy enough, and would seem to be nearly universally

agreed upon as an appropriate purpose of schooling. Yet considerable variations in the definitions of the basic terms, "right" and "wrong" form the horns of a dilemma which bedevils the ordinary classroom teacher's practice. The problem of how to go about forming character so that it conforms to a generally accepted notion of goodness without being so particularized as to stir angry debate or controversy within the community (of classroom, school, or town) is a complex one. Nonetheless, from the start schools were expected to develop character in their students, and this is still true today. What exactly is meant by "character" is not always clear, yet it is clearly something most people recognize as important.

What is Character?

Character training generates a constant stream of debate and discussion in part because the concept of character is another of those ideas which is fixed in everyone's mind, but in a slightly different way for every person. Perusal of the literature surrounding character training reveals this, as each author presents his or her unique approach to character education without presenting an explicit definition of character, even as some other author is picking that implicit definition apart in an attempt to show that the method advocated by its holder is flawed from the outset. Davis (2003) notes the evasive nature of the term character, complaining that "I expected that most discussions of character education would begin with a definition of 'character,' [but] I found few that did. . . . Character education does not seem to be a subject that invites clarity. Nor am I the first critic to notice the rarity of definition" (p. 52).

Davis himself struggles with a definition, beginning with the rather vague statement that, "By 'character' I mean the relatively settled general disposition of a person to do

what is morally good" and offering "rough synonyms" for the term: "good character;" "moral character;" "moral integrity;" and "virtue." Arguing that "character can be analyzed into a set of 'traits,'" Davis generates the very list of virtues typical of old-fashioned character education advocates: honesty, compassion, perseverance, respect for others, courtesy, and so on.

While this example is presented here in order to demonstrate just how difficult it is to arrive at a clear definition of "character," Davis makes an important point about the way in which we typically use the term. "Character" is used in a way that seems to require an understanding that if we have "character," it is good. Thus, character education assumes that the virtues it inculcates are applied for the good; diligence, fortitude, perseverance, and determination in pursuing evil are no longer virtues. Yet, taught apart from the kind of discussion that makes the distinction clear, one could ask, as Davis does, "What's wrong with character education?" Indeed, many of the traits listed in state-mandated courses for character education can be developed in a way that is not moral. Using the means (a list of virtues) to define the ends (what constitutes good character) of moral education is not merely circular; it may be self-defeating.

Rivers (2004) defines character as "personality traits that are privileged" (p. 250). Attempting to clarify, he goes on to say that " 'Character' is a term used to refer to a composite of personality traits that are essential.. . . 'personality' . . . doesn't suggest inner effort or even choice. . . . we do not choose our personality; we do choose our character" (p. 251). This rather poor attempt at defining character at least highlights a common notion of the character education thinkers: character can be taught and molded, because it can be chosen.

In an attempt to establish "the science of character education," Berkowitz (2002) defines character as

> "an individual's set of psychological characteristics that affect that person's ability and inclination to function morally. . . . character is a complex psychological concept. It entails the capacity to think about right and wrong, experience moral emotion (guilt, empathy, compassion), engage in moral behaviors (sharing, donating to charity, telling the truth), believe in moral goods, demonstrate an enduring tendency to act with honesty, altruism, responsibility, and other characteristics that support moral functioning. . . . I attempt to redefine character as a complex constellation of psychological dimensions of a person" (pp. 48-49).

This "redefinition of character" does not serve to clear up the already murky waters surrounding the concept, however, and Berkowitz' own admission, a few sentences later, that, "I am not wedded to this particular definition" (p. 49), seems to undermine its purported scientific value.

Because most people accept the list of virtues typically presented as those belonging to good character as worthy traits to be developed in children and youth, and because such a list serves as a useful means of identifying places in instruction where issues of character development may be being addressed, it is worth noting that the schools in New York, the site of this study, were required by legislation adopted in 2001 to provide

> "Instruction in civility, citizenship and character education. The regents shall ensure that the course of instruction in grades kindergarten through twelve includes a component on civility, citizenship and character

education. Such component shall instruct students on the principles of honesty, tolerance, personal responsibility, respect for others, observance of laws and rules, courtesy, dignity and other traits which will enhance the quality of their experiences in, and contributions to, the community. The regents shall determine how to incorporate such component in existing curricula and the commissioner shall promulgate any regulations needed to carry out such determination of the regents" (Kadamus, 2001).

The list of virtues contained in the New York State legislation is a typical one, and leaves room for additions ("other traits") that may seem appropriate to particular circumstances, schools, or school districts.

Often the shaping of character is couched in terms of citizen education: "Character education, because it deals with relations between and among individuals and groups, conditions of civil society, and significant public issues, is central to citizenship education" (Howard et al., p.189). Thus, one would expect the social studies teachers to carry the burden of character education. Yet English class has traditionally been seen as an appropriate place for training in character, because the contents often provide convenient vehicles for discussion of moral issues.

English Classes and Character Development

The supervisor of Boston Schools, G. H. Martin, described English literature as "the unseen force in character-making" in 1899 at the annual meeting of the National Education Association (Journal of Proceedings, NEA). William Torrey Harris included literature and art as one of his "five windows of the soul;" and in 1905, Reuben Halleck presented his views of "The Value of English Literature in Ethical Training" at the forty-

ninth annual meeting of the NEA (Journal of Proceedings, NEA). This line of argument

remains as strong today as it was over a century ago; arguing that "character education

has become the fastest growing school reform movement in the United States" (p.113),

William Edgington (2002) promotes the use of literature for children and adolescents for

advancing character education. According to Martinson (2003), "stimulating the moral

imagination" is an important classroom activity, and the usefulness of literature for

presenting moral issues is evident to most classroom teachers (Gilness, 2003).

Despite—or perhaps because of—the constant discussion of the topic of character

education, there are nearly as many ideas about what it is as there are about how to do it.

Everyone has some sense of the meaning of the term, "character education," and

definitions range from the apparently very simple to the rather complex. One can accept

nearly all of these definitions on their face, but considering how they play out in

instruction makes some more unwieldy than others. At the simple end, many accept

Lickona's (1998) definition of character education as "the deliberate effort to cultivate

virtue" (p.78). This leads to a listing of the virtues and debate over what traits are virtues,

which ones can be taught, whether a virtue is a constant good in itself or dependent on

situation, and so on. Soon, the definition itself seems inadequate. Authors of books and

articles on character education (Gibbs and Earley,1994; Bennett, 1993; Lickona, 1991;

Wynne and Ryan, 1997) present varied lists of virtues or values such as courage, honesty,

compassion, respect, perseverance, responsibility, and so on; all of these, being moral

qualities, are open to interpretation and have wide variations in definition. Yet, despite

the difficulty in arriving at universally agreed-upon definitions of the specific traits

involved, or for the overall concept, nearly everyone agrees that schools must, can, and

do address the moral development of the pupils in terms of character development. Dewey's (1909) observation that teachers teach morals "every moment of the day, five days a week" (p.3) is just as true today as it was a century ago. Much of this instruction in morals is part of the "hidden curriculum" (Giroux & Purpel, 1983), yet it emerges more prominently in English class, where literature serves to raise issues about virtues, values, and visions for a just society.

Three Approaches to Character Development

Attempts to make known the possible ways of approaching character development in schools often begin with classifications of various schools of thought regarding moral training and citizenship. In an interesting analysis of the major philosophical traditions of character education, Howard et al. begin with Lickona's (1989) assertion that character has three parts: "moral knowing, moral feeling, and moral behavior. Good character consists of knowing the good, desiring the good, and doing the good" (Lickona, p. 59, in Howard et al., p. 190) Howard et al. point out that these three parts receive different emphasis from each of the three major approaches to character training.

> ". . .there are three major approaches to character education: the cognitive-developmental approach (often called moral education) gives primacy to 'knowing the good,' the caring approach emphasizes 'desiring the good,' and traditional character education, which sees 'doing the good' as fundamental" (p. 190).

Doing the Good: Traditional Character Education

The three traditions of character training all require all three parts, but put the emphasis more heavily on one or another. In the early days of America, traditional

character education was based on Puritan principles of virtue, then on mainstream Protestant ones, and then on "standard middle class" ones (see McClellan, 1999, pp. 45-55). Americans are doers; hewing a nation out of the wilderness and building a country that stretched across a continent meant that from their arrival on the shores of the New World, Americans had much to do and valued hard work, determination, and cooperation. Perhaps this explains why the tradition of character education emphasizes doing the good over knowing the good and desiring the good. These other features of character development are certainly required for doing the good, but, in traditional character education are received with little debate from those (preachers, teachers, elders, parents) who already know what is good and desirable—reflection on the basic virtues and their meanings is not necessary, because it is already established; the important part of character education is the development of action and habit, so that doing the good is a kind of conditioned response.

Howard et al. call this approach an "Aristotelian tradition." This approach creates "a community environment that imbues youth with the virtues and then reinforces them through formal instruction, visuals . . . positive peer culture, and ceremonies" (p. 191). McClellan (1999) points out that this approach "showed little tolerance for cultural diversity" and played "an important role in eliminating the differences that set immigrants off from the mainstream American life" (p. 55) during its heyday from the late eighteenth century until the turn of the twentieth century. Essentially, this kind of character training passes on an already existing view of virtue and values with little room for argument. Bennett (1993), Kilpatrick (1992), and Ryan and Bohlin (1999) are cited by Howard et al. as contemporary examples of this approach. The rebound of the

approach in recent years may be partly a result of fears generated by incidents like the Columbine school shootings and a negative reaction to the multicultural sensitivity debate of the 1980's.

Knowing the Good: A Developmental Approach to Moral Training

The second school or tradition in approaching moral training was "moral education" which "had progressive change as the primary goal of schooling and developmental—process-oriented—pedagogy for character education" (Howard et al., p. 192). Labeling this approach "Socratic," Howard et al. point out that it emphasizes reasoning, and teaches "children to engage in critical thinking and to have a process on which to call in making decisions and actions" (p. 192). Based partly on progressive political premises, this second version of character training shifted the emphasis from personal moral and private conduct to social behavior aimed at the betterment of society, and aimed to train students to "engage in the skills of democratic citizenship: deliberation, problem solving, and participation in governance of the group" (Howard et al., p. 192). John Dewey (1909; 1916) was a major figure in promoting this version of character development as was Frank Chapman Sharp (1917, 1927); Howard et al. cite Piaget (1932) and Kohlberg (1981; 1984) among those whose "developmental approach" to moral training is rooted in this second approach to character training.

The shift away from character education toward moral education is attributed by McClellan (1999) in part to the shift toward a greater distinction between private and public behavior and "a fear that character education was, or would be seen as, improperly invading the privacy of students and families" (in Howard et al., p. 194). In addition, Howard et al. point out that the 1960's brought an erosion of moral authority with

opposition to the war in Vietnam, the growth of the civil rights movement, and the "sexual revolution;" these undermined the received nature of traditional character education. Citing the Supreme Court's decision in Engle vs. Vitale (1962), Howard et al. argue that "there was growing pressure for separation of church and state. . . for the many who see religion as the foundation of ethics, banning [teacher-led] prayer was perceived as tantamount to banning character education" (p. 194).

Even as character education took a back seat in character development efforts, ways to replace it were coming onto the scene. Values clarification (Raths, Harmin. & Simon, 1966; Simon, Howe, & Kirschenbaum, 1972) proposed an examination of one's values using exercises to "increase personal awareness of and/or make critical decisions about the values they held" (Howard et al., p. 194). However, the inability of values clarification to make distinctions between ethical values and other kinds of values, like preferences, tastes, and likes or dislikes, meant that it never had the influence of the moral education approaches that favored a cognitive-developmental approach, typified in Kohlberg's philosophy of moral development.

Kohlberg's method used hypothetical ethical dilemmas to encourage students to develop critical thinking skills and come to their own decisions about such dilemmas. Based on a theory of developmental stages in moral growth, the approach claimed the existence of "universal and invariant stages of development" in moral reasoning (Howard et al., p. 195). While this model shifted the focus of moral training from received notions of virtues to the development of ways of thinking about ethical issues, this central claim of stages of development led to criticism (Simpson, 1974; Sullivan, 1977; Shweder,

1982), and eventually shifted the focus of moral training from "knowing the good" to "desiring the good," i.e., caring about others.

Desiring the Good: The Caring Approach to Moral Training

Feminist theorists attacked Kohlberg partly on the grounds that his original samples for developing his theory were entirely male and mostly white; this meant they were not representative of the development of girls and women, and perhaps not even of males who were not white. Gilligan (1977) explored the moral development of girls and argued that Kohlberg's model was inadequate; she proposed a psychology of moral development with a focus on care. Noddings (1984; 1992; 2002) took up the same theme in philosophical terms, and a third approach to moral training moved the focus toward questions of relations with others.

> "The caring approach to morality differs from the character and developmental approaches in significant ways: (a) morality of care is relational rather than individual; (b) it gives primacy to moral emotions and sentiments, claiming these to be the stimulus to moral action and moral reasoning (not always in that order); and (c) care does not require that moral decisions need to be 'universalized' to be justified. Creating and maintaining relationships. . . is at the core of the caring approach's prescription for schools" (Howard et al., p. 195).

The idea that creating and maintaining relationships is at the core of this approach to moral training has caused many thinkers to reflect more seriously on the nature of the community that is created within a school. Sizer and Sizer (1999) muse on the importance of what actually happens in school as opposed to what rules are promulgated

for the moral edification and advancement of the students. "They watch us all the time," they observe, "they listen to us, sometimes. They learn from all that watching and listening'' (p. xvii). Perhaps the most important contribution to moral training made by the caring approach is the realization that schools must model what they teach, and that moral growth and development does not occur in a vacuum.

Other Models of Moral Education

While the three approaches to moral training described above capture the chief differences in perspective that guide them, there are many other ways of looking at moral education, each with its own slant on what is most important. However, no advocate of any one program or approach claims that it should exclude elements of the others, and so the result is, ironically, rather like the typical approach to grading: a hodge podge of elements eclectically combined by the individual teacher aimed at creating an approach with which he or she is comfortable to make judgments, express opinions, and build personal relationships with his or her pupils.

Some of the alternative ways of approaching moral education are summed up by Joseph and Efron (2005) as "seven worlds of moral education." Admitting that each "world" is based on its own assumptions of best practice, learners, and morality, Joseph and Efron concede that none of the approaches "exists in isolation; nor are their purposes . . . opposed; they may, in fact, share several characteristics" (p. 525). A brief description (drawn from Joseph and Efron, pp. 525-531) of each of the seven approaches will serve to indicate their salient features:

1. Character education: Based on a list of virtues, this approach strives to inculcate those virtues in the young. Typically, the list of virtues includes self-discipline, compassion, responsibility, and so on, as discussed above.

2. Cultural Heritage: While typical character education programs draw their list of virtues from the dominant culture, cultural heritage models draw their values from traditions of non-dominant cultures. Afro-centric and Native American schools that teach the language, customs, and history of a non-dominant culture may follow this model.

3. Caring Community: emphasizes an ethic of care, i.e., nurturing, closeness, emotional attachment, and respectful, mutually supportive relationships.

4. Peace Education: extends the caring community beyond the school; emphasizes harmony and the "well-being of humanity;" includes conflict-resolution, human rights education, global education, environmental education, and peace studies.

5. Social Action: Guided by the concepts of justice and compassion, this approach is focused on the political nature of society. Students are seen as social agents who can effect change by critically examining unjust situations and correcting them through political action.

6. Just Community: Based on Kohlberg; the goal of moral education is the movement of students' development from lower to higher stages of moral reasoning. Students, teachers, and administrators discuss and address matters of mutual concern, so that "students clarify and refine their thoughts while listening and responding to other points of view' (p. 530).

7. Ethical Inquiry: "Students engage in 'moral conversation' centered on dilemmas. . . this. . . approach is grounded on the premise that deliberation promotes students' moral development…teachers invite students to investigate values or actions and to imagine alternatives. . . .Teachers guide discussions on the moral dilemmas embedded within subjects across the curriculum" (p. 531).

What emerges even from this brief look at other approaches to moral training is the wide variation in understanding of what is most important and worthwhile in such undertakings. Every community, every school, and every classroom has multiple currents of morality swirling through it. That teachers hesitate to step into this river is natural enough; that they dive in head first and steer their course through these currents is equally natural to them and is typically seen as part and parcel of the occupation of English teacher.

An Environment for Moral Training

We have already seen that the care approach to moral training calls for an environment which fosters such development. It is natural that each method would be best served by the creation of an environment that would encourage the practice of such virtues as were believed to be important for good character. Most American public schools, in fact are already set up in such a way as to foster traditional kinds of character education. This fact becomes clear in Wynne's (1997) exposition of "for-character education."

In a section entitled, "designing appropriate environments for for-character education," Wynne presents four example virtues and ways to promote (or discourage) them. The activities, because of their very traditional nature, are revealing. To

encourage diligence, homework and well-paced instruction and feedback are called for; to encourage obedience, written rules for the classroom and steady and clear enforcement of school rules are needed. Cooperativeness can be fostered by extracurricular activities, team sports, service clubs, cooperative learning activities, and so on. And loyalty can be encouraged through a daily salute to the flag, carefully planned assemblies for "appropriate causes," a school pledge, and pep rallies (pp. 68-69). While these things have a kind of retrograde ring to them, Wynne argues that careful implementation of such activities will develop good character, since "it is indisputable that the character of persons reared in different environments are [sic] affected by such environments. . . . Environments affect the modal character of the people born and living in them, and also help clarify the inhabitant's definition of 'good character'" (p. 74).

Wynne's argument is at least implicitly shared by other approaches to moral training; each of the seven "worlds of moral education" introduced above requires a unique environment consciously formed to foster the goals of the approach. Some are sufficiently specialized (e.g., the cultural heritage approach) as to be unlikely to succeed in most places because of pre-requisite qualities on the part of the teachers (e.g., shared cultural background or "deep understanding of the students' culture"). The approaches that have underlying themes more directly related to the "mainstream" of American public schools, and can draw upon elements commonly found in public schools for the basis of the required environment, are more likely to yield results.

CHARACTER DEVELOPMENT, ENGLISH CLASSES, AND GRADING AT MELLMAX HIGH SCHOOL

Over the course of the year's interviews, all of the English teachers at Mellmax High School spoke of moments where they found themselves teaching "life lessons."

Sometimes these sprang from discussion of a short story or novel. At other times, they resulted from an essay question or journal entry that required a conference. At still other times, they came from a frank personal conversation with a student about his poor conduct or disappointing effort. Clearly, the moral implications of these lessons were clear to the teachers from the mere fact that they labeled them "life lessons" as opposed to ordinary day-to-day lessons. The label implied something even more significant: these lessons were more important than lessons on parallel structure or subject-verb agreement. They were lessons that would help shape the student's character.

One would expect that if such lessons were so important, they would somehow make their way into the grades of the now-improved student. Yet the teachers invariably denied that they allowed such things to influence their grading, at least initially. As the interviews continued over the course of the year, and the teachers reflected more and more on their grading practices and the ways they judged the more subjective parts of their students' work, they began to admit that there was at least a possibility that the way they saw their students developing as people did in fact play out in some measure in their grades.

Literature as a Vehicle for Discussing the Moral

Like most English teachers in America, those at Mellmax High believe that literature is an excellent vehicle for discussing moral issues. Common literary terms lend themselves to thinking in moral terms: conflict, internal or external; protagonist; antagonist; hero; heroine; and so on. The novels read in class become natural sources of discussion about courses of action, decisions to be made, and ways of treating one's neighbor. The link to the world beyond the novel is almost immediate. As Mary Minton

observed, "I think through literature there is obviously a lot of room for discussion of events that happen in a book, and I think it's probably one of the areas where you can look at morality. But I think you have to be really careful because you can't impose your feelings about that. I think that morality is obviously morality. . . . I can remember times when we were reading *The Diary of Anne Frank* when kids would talk about the Nazis, and that gave you opportunities to explore things. I think that morality has certainly declined in school. But I think as an English teacher you have a lot of times when you can really get into some good discussions and I think it's difficult, but I think it's important."

Karen Kistner suggested that she usually discussed morality in terms of the fictional characters her class was reading about. She seemed to avoid drawing upon her students' own experiences, but did if an issue arose directly in the classroom. "We get to talk about moral development from the characters we are discussing in their literature—we get to talk about the moral development of these characters. Do we get to talk about their [the students'] moral development? Once in a while something happens in the class and you get to talk about cheating and lying, being kind to your neighbor, this kind of thing. Maybe you have a chance to discuss that in class. But essentially it's done with literature, with characters in literature."

Jeanette Jones was more eloquent. "I think English is a reflection of life. Statements on life. We study war; you study human development; you study man's inhumanity to man. You really talk about all those things with kids, and it's a great impact. . . . I think things they read reach them on a personal level, and then they write about it. . . It's not cut and dried like math; it's not necessarily content oriented." Jeanette remarked on the

way one of her required novels, *A Thanksgiving Treasure*, affected her junior high students. "They love that Thanksgiving book. That's what we were talking about today. There's something in that book that reaches them. I don't know what it is; it's just a wholesome book, and there's some message in that book, that it just sort of sets the tone for the year."

Hester Hypoint admitted that she saw part of her role as the moral development of her students, and that this development was fostered by literature. "At times. In conjunction with literature especially. Things will come up, and then you get off in other areas. And …yes, you do. Yes, I do [see part of my role as advancing the moral development of the students]."

Frank Fender was more specific, and suggested that an element of personal connection was important to developing character. "I think [English] is probably one of the most important pieces of education. Character? Well, there too, I mean, you really can't tie in a personal conflict in a math problem or science, but English offers you every opportunity where you can make a relation to them [the students], where you can make a lesson, a life lesson. I think that's my strongest part of being a teacher."

Although David Dutcher heavily emphasized intellectual development and reasoning skills in his English classes, he also noted the value of literature in discussing moral issues. "If you can enlighten a student about a theme that promotes morality, then yes, you are affecting that student in a moral sense." Referring to *The Scarlet Letter*, David explained how the scene in the novel where Hester Prynne removes the letter and her daughter, Pearl, will not approach her allowed him to "preach a little bit" and explain that "you cannot just cast aside the consequences of your moral wrong-doings. More than

192

that, you can't simply change your identity by simply throwing away those things that have made you what you are."

Talking about character education and values, David complained that morality in American public schools seems to be addressed in terms that avoid the use of the term "morals," as if this were an imposition of religion or doctrine. This highlights the sensitive nature of forming character in schools; teachers worry that their discussions of moral issues will be seen as proselytizing, and are sometimes unsure "where the lines are;" among the veterans, a certain sense of caution (expressed in Mary's warning to be careful, Karen's apparent unwillingness to directly engage the students in personal moral discussions, and Hester's almost vague response concerning the role of literature in moral development) was evident, perhaps because of years of experience in dealing with sensitive issues—and a small share of conflict over the years.

Alice Andrews expressed the combination of optimism and disappointment in dealing with issues of broad moral significance that seems typical of the novice teachers. "I hope that they leave my class having learned some kind of moral lesson. That's the good thing with English, being able to address so many issues like that. Many times, though, I walk away and feel frustrated because I wanted a greater appreciation of something, or I wanted them to be a little more honest. …If I could just change one student's idea about something…I think that's the great thing about being an English teacher. We have such a broad range of things we can discuss with them, and we can be more personal than the other teachers."

Following the pattern of the literature review for this chapter, we will next consider how the three major approaches to moral training found expression in the practice of the Mellmax High English department.

Doing the Good at Mellmax High

The character education approach, which emphasizes doing the right thing as the central part of moral training, translates into common practice in two ways: inculcation of a list of desirable virtues (this is discussed below in greater detail) and the use of school activities and assignments to encourage and develop the habit of doing the good. All the teachers remarked at one point or another on ways that behavior affected classroom outcomes. Effort especially was seen as a kind of habit of virtue; in the previous chapter, effort was examined in terms of its value as a sign of ability and source of success. Here, it is seen as a moral necessity tied to life: learning to apply oneself is important, with or without success, because living requires constant effort.

Remarking on the comments sent home for both the report cards and the progress reports, Mary Minton observed that things "like not doing your homework would affect your grades, of course.. . . I would imagine that comments about things like not doing homework or misbehaving in class or not taking opportunities to do work over that they didn't do well on would affect their grades. That may be behavioral, but it's going to translate into a grade. I wanted to have an effort grade. I think because we're human beings, effort is important. And we used to have one. I remember years ago when we did have one, I would get more calls from parents about the effort grade than their average." This last remark suggests that parents, too, expect their children to learn the importance of effort, and, for many, this lesson outweighs their academic achievement.

Learning how to behave in class was another important part of doing the good, in the eyes of Frank Fender. "I think after five weeks you can determine behavior, their commitment, whether they're going to do their work or not. The grades are secondary to behavior. . . . if a student is a disruption, if they're talking while I'm talking, if they're talking about everything that's not related to what the English class is about, I take that into account. Is the student late, do I have to address the student a number of times? Is the lack of homework a constant, where it's a pain for me?

"If a student is late all the time, they're missing some of my class, and I don't think it's my job to repeat myself if they're late without an excuse. If they don't want to make up their work because they're absent without an excuse or they skip my class, that's not my job; it's their responsibility to make up the work. So, and if a student is a behavior problem they're supposed to be listening to me, they're supposed to be doing their work, so it does affect their grades."

Asked how he balanced behavior in his grading, Frank explained that "usually the bad behavior coincides with the poor grades, I mean the poor performance. Very rarely do I have a student that behaves poorly in class and still excels. . . . If a student has like an 89, and has been a great student for me, you know, then I'll probably bump it up to a 90. If the student has been a problem, then there's not a chance."

Catherine Carney had similar concerns; doing the good meant acting like interested pupils. "By the end of the first ten weeks, I'm assessing the students themselves. Are they listening, are they working, are they paying attention? Who's developing good work habits? And I know which ones are really trying and which ones aren't."

Gail Goodwin remarked that while her numerical grades were not, she hoped, influenced by her perceptions of her students' characters, she used the comments on the progress reports and report cards as rewards and punishments of character development. "If a student is a good example in class, I find a comment that says that. But if she is uncooperative or hostile, I look for a comment to send that message home."

All the teachers, veteran, established, or novice alike, agreed that developing the habits of behavior expected of good students was a fundamental step in their character development. Betty Browne summed up the typical stance. "I don't want them to think they can just throw something sloppily together. I conference with them and have them do it over. I kinda like to reward them to do better, so yeah, I give them more for doing it again. We'll discuss what we think is fair. There are bright students who don't work up to the level they should, and so I'm not going to just give them the grade they're used to.

"The student that I get worried about is very negative, distracting, and talking to everybody. Well, that's OK, we'll talk about it. I try to handle things by myself before I take it to another level. But if I'm constantly pulling them aside, pulling them aside, that's not satisfactory, so I'm going to have to take another step. The student that just turns off in the classroom, just zones out, and has no desire to be there—they're just clueless—that's not satisfactory, because they just don't have that work ethic, and that's going to be a problem."

"They do group work, and there are times when they have to be working independently. There are ways to behave in either situation, and they need to learn how that should be."

Knowing the Good at Mellmax High

Moral education, the second major current in moral training, differs from character education in that it emphasizes the development of skills in thinking through moral issues and places a premium on personal decision-making about issues of right and wrong. Unlike character education, which begins with a received list of right behaviors, moral education begins by thinking about right and wrong, and what is the right thing to do at the moment, based on one's way of thinking. The teachers at Mellmax High did not seem to favor one approach over the other, but used parts of each as they saw fit.

"I think I asked them questions about the rightness or wrongness of certain things," reflected Mary Minton. "But I always made sure—I always tried to be very objective, and even though I may have had a side in how I felt, the only thing I was always looking for was support; I thought that was important. 'You don't have to agree with me, but you have to be able to support how you feel.' But I do think I gave assignments where they were able to explore that and use their own feelings and ideas and experiences. But I think you have to be careful when you're grading to be able to still be objective because some students have some very different ideas than you may have."

Karen Kistner, still trying whenever possible to limit moral issues to the stories being read, was nonetheless able to find ways to encourage moral thinking. "Sometimes your writing assignments can be based on whether or not you thought this man was morally right in what he did. . . we discuss that. I don't know if we have as many writing assignments on that, but it would only creep into grading if they didn't write the assignment." (This last remark is interesting for its hybridization of both approaches in one assignment—while the student is to be thinking through the moral processes in the

character's behavior, his own behavior is being guided by the rule that all good students *will* write this essay!)

Luann Lender presented the moral education approach in an elegant declaration of her role as an English teacher. "I like to think I am giving them not only the mechanical skills of punctuation, grammar, and you know, being able to write a complete sentence, but a way of looking at the world and at life, and approaching the world and life. I like to think that I was teaching them techniques of dealing with things in life, and helping them see that you don't have to be overwhelmed by things, that there's a way to approach whatever is in front of you to deal with it. So I like to think that I didn't just teach them *Animal Farm,* but I taught them a way of looking at any kind of literature or any kind of situation and learning how to interpret it and react to it and react with it. So I don't think I was just a mechanics instructor. I'd like to think that I was a lot more than that."

In an indirect lesson on knowing the good, David Dutcher described his reaction to a display of poor character versus poor performance. "If a student writes something and is very insincere, then I think the student is really leading you into a situation where you're making a personal judgment call. If a student does a poor job, but there is a serious tone that is indicative that the student is conscious of the audience and the purpose, you can try to be as objective as possible in your comments and the grading without making it an indictment on a personal level." Describing incidents where unprepared students wrote what David interpreted as "facetious" answers, he explained that he wrote comments "to the effect that 'I know that you're intelligent, and I don't appreciate it when you degrade yourself with responses like this.' And it seems to have done the trick, because I haven't seen that as often…just a couple of incidents." Learning to know the good is more

challenging, perhaps, than learning to do the good, and involves both the teacher and the student in making more personal judgments about situations and themselves.

Alice Andrews confirms the point**:** "it takes more time to sit there and really think and make connections, either to their own lives, to other pieces of literature, to underlying ideas, so I …weigh that more heavily because it's more personal, and it's me being able to understand how they understood a point or how they're feeling, rather than just regurgitating information."

Desiring the Good at Mellmax High

The third basic approach to moral training, the caring approach, posits a personal relationship of a caring nature between the teacher and the students and between the student and other students. This, too, was part of the eclectic approach of the teachers at Mellmax High.

"Acceptable behavior to me was showing respect for me and the other students," declared Mary Minton. "Did behavior influence grading? I think if a student is constantly misbehaving, I think their bad behavior can, and it can hurt the other students. . . .I don't know if it influences their grade, but there's an interconnectedness there. Was I fairly grading the ones who were acting out? I think I tried to work very hard with those students that misbehaved."

Luann Lender asserted that she "would be ashamed if I had to say I let my opinion of a student affect their grade. I don't think of students as "good kids" or "bad kids." This is a person, this is not a good kid or a bad kid."

On the other hand, Luann saw writing as a deeply personal undertaking, and felt she could not separate the students' identities from their papers. "I don't think you should. Writing and interpretation is too personal a thing to separate it from who they are, because when you interpret a piece of literature, you're exposing yourself and your own experiences. I don't see how you could separate the two. This isn't math, where two and two is always four. This is English, and interpretation, where two and two might be six and a half, depending on what you bring to it and how you interpret it."

Ellen Enders adjusts her idea of "satisfactory behavior" for each child as she gets to know her pupils. "If I'm grading a group activity where they have to work with others, then of course their behavior affects how I grade them." She tries to conference over grades in order to impress personal feelings about performance on the students and encourage them to work to their potential.

"Grades aren't a reflection of the person; you're not a bad person because you didn't get 100. And that's what I hate about the Regents' exam; they shouldn't all have to pass it to prove they're worthy citizens in society. On the other hand, I've got kids right now who do nothing and are capable, and that's not fair to the other kids. At the very end of the year, if I have a kid that's been really good for me all year and he has a sixty-three, I'll give him the two points. But I've never taken points away from another kid who's been a problem for me."

Asked if "good kids get graded differently from bad kids," David Dutcher replied, "Well, the 'good' with the quotes around it, I probably have a mental idea of what that means better than 'bad kids.' The reason I say that is because there's a very broad category with a very wide-ranging spectrum for what I guess you would call the opposite

of 'good kids.' I don't want to call them 'bad kids.'" When asked if he separated the students' identities from their work, David argued that this was not possible, and that "I use a rubric" was "the easy answer," admitting that there was some element of subjective judgment in his grading, as heavily dependent on the use of rubrics as it was.

Jeanette Jones expressed a different level of concern about her junior high students. "In 7th grade, you worry about transition to the building. Is this child gonna be this way all year? Or is it just because he's made a poor judgment? I've found that the comments on attendance bring more comments from the parents than the actual grades." Again, this reaction from some of the parents suggests that they, too, believe that certain kinds of moral behaviors can prove to be more important than academic achievement.

Jeanette also believed that the English classroom needed to foster confidence in the students. "I think that writing is a very personal thing, and if they don't trust you, they aren't going to write anything. And I think some of our kids are so frustrated so easily and negative about life. That's why they won't write. They don't want to let anything out on paper. And I've had kids really spill their guts, you know, even with these awful writing assignments that we give now, you know, and that's your role. You're not just a teacher, you ARE a social worker. You ARE a parent, you know, you are all those things." Her approach seemed to get results. "I work so hard with these kids, in bringing them in and talking to them, that usually I don't have that many that are outstanding behavior problems. You know, every year you might have one or two."

Jeanette's moral concerns were not limited to behavior. "I worry about fairness now because I don't think the [NYS eighth grade] test is fair. I don't think a lot of what I'm doing in 8th grade is fair. I think it's beyond them and it's frustrating them, you

know…and what do you do about that? The first year of that test, you know, we worked so hard and worried so much about what they were gonna get, and I realized, 'they're gonna get what they're gonna get. You might as well not make them totally miserable.'"

Learning Virtue

While character education was not the only strand of moral training evident at Mellmax High, it was an easily detected one, and the use of the idea of the virtues, although they were not labeled as such, was evident in the practice of the English teachers.

Surprisingly, Gail Goodwin actually rattled off a list of virtues she hoped to develop in her students: "I hope I make a good impact on them. A lot of what I do in my room is trying to get them to develop organizational skills, a work ethic, pride in their work, responsibility, you know. I also do a lot about character development and self esteem. So I'm hoping in the big scheme of things, I'm getting across what I want to." Explaining that she used a set of quotations on character traits such as perseverance, hard work, and honesty, and aligned these with the current reading assignments, Gail acknowledged that connections made between the things said in class about character and the essays students wrote for assignments were rewarded. "Yeah, I guess in a way it works into the grading."

Other teachers clearly valued specific virtues; diligence was one of these. Jeanette Jones argued that steady application and endurance were worth reward.

"Yeah, oh, sure, and it also might affect, you know, if you said, 'Well, this person is trying so hard and has a sixty-three, you might not give them a sixty-three [i.e., a failing grade]. But if somebody really goes out of his way to give you a hard time, you might *not* give them a sixty-five [i.e., a passing grade]."

Hester Hypoint agreed that diligence was important. Asked if she could see progress in her students, she said, "In some. In quite a few, actually. Some not, mainly because they're not doing work, and refuse to do it. And this is where the problem lies. In fact, that's where a lot of the problem lies. In not doing. Where they're taking zeroes. That's where it is. Many of them are weak, but they could squeak the 65 if they did the work."

Another valued virtue was honesty. Nearly all the teachers lamented the general trend toward lying—about homework, about effort, about understanding. Catherine Carney admonished her students at the beginning of the year.

"I expect them to be telling me the truth," she said, "and I tell them so. How can I really help them if they aren't honest with me?"

Obedience was another highly prized virtue, although it was not called by its name, perhaps because it seems too suggestive of submission to a stronger power. Instead, it is called "following the rules." Explaining how she assessed satisfactory and unsatisfactory behavior, Hester Hypoint put it very simply: "If they follow the rules, it's satisfactory; if they don't, it's unsatisfactory." She did not, however, believe that she allowed her assessment of a student's obedience to affect her grading practices.

"I've given 90's to kids whose behavior was awful, and I've given failing grades to kids whose behavior was awful. …I don't take a lot of personal things into account. If the test is objective, well, these are the numbers; and if it's an essay, it's got to fit the rubric and that's how it's graded." One could argue, however, that rubrics are often little more than a list of rules for completing an assignment. If one fails to follow the rules out of inability, this is ignorance; if one fails to follow them out of contempt, this is disobedience. And failing to follow the rules, then, does indeed affect one's grade.

Respect is a virtue already mentioned in regard to the caring approach to character development; respect for teachers, classmates, and oneself is essential to moral development. Several of the teachers remarked upon the importance of mutual respect in fostering moral development. Ellen Enders described her rubrics for group activities; like those of many of the other teachers, these included a category that could be used for rewarding or punishing behavior and judging cooperation with, and respect for, classmates.

A final representative virtue was responsibility, a standard on almost everyone's list of virtues relevant to schooling. In recent years, this has also been called "accountability," although one might argue that the two are slightly different, the latter implying exercise of some kind of authority that requires the former, whereas responsibility is a broader concept that applies to a sense of duty or moral rectitude. The Mellmax teachers did not engage in such hair-splitting, however, and used the terms interchangeably. Giving an example of lack of responsibility and the consequences thereof, Jeanette Jones told of an unfortunate student.

"He knew for 3 weeks that he had this test coming up, and he didn't read it [the assigned novel]. So he has a zero. They're promoted without doing any work. Nobody repeated 6th grade this year, not one student out of 150. Now, that's not right. And then they come here, and when they're held accountable, it causes us all kinds of stress. They come here and they don't do any work and they dig themselves into a hole, then the parents blame us—'they were fine until they came to the junior high.'

"In the end sometimes, if you have a student who's very lazy, that grade is going to reflect that. But I think that it's not a personal thing, if that makes any sense. He didn't

do any work, so that's what he earned. So I try not to make a judgment, and say, you know, 'You're stupid.' I might say, 'Well, you didn't do the work, it's not that you're stupid. But you didn't do anything, and your grade has shown that.'"

Using Grades to Encourage Habits of Virtue

While most of the English teachers at Mellmax High insisted that they kept moral training quite apart from their grading, some used grades to foster—or perhaps coerce— habits of virtue.

Catherine Carney uses zeroes as a means to keep even the better students working; if assignments are missing, they are calculated as zeroes. "Today, I gave them time to work on a homework assignment at the end of class. To me, if that paragraph comes in like they did it in homeroom, I probably will remember that they were not doing anything in class. Effort and behavior to me are almost the same thing. Yeah, they must be. I see effort in class….If I don't see them working in class, I do start to label a student."

David Dutcher assigns many literary essays; these are bound to have room for subjective judgments, despite the highly specific rubric (which, again, calls for following rules). "My policy with a disappointing writing assignment is to require a revision. But I grade the first draft as a safety so the student can't say I wouldn't accept it. I replace the first grade with the second grade." Since these are poor papers or ones that do not show what is expected, this policy shapes character, whether David sees this or not; the poorer student must try again, and be diligent; the better student must try harder, and act in keeping with his ability. Better papers get fewer comments; weaker papers or papers that do not live up to the expected potential of the student get more comments; often these are exhortations to persevere.

Like Catherine Carney, Alice Andrews was frank about the way she deducted points to punish the indolent. "Students who are never in class or send a message that they don't care do get points off the subjective parts of tests to send a message."

Karen Kistner remarked, "I get many cases where students have really created negative rapport between the two of us. I try to the best of my ability not to let that interfere with my scoring. I really don't think it does. I think the objective part of the test is enough to at least adjust that for me. If I have a negative rapport with a student and he has a 90 on the objective part of the test, I don't look at places to take off points."

While Karen was scrupulous about not taking points off for "negative rapport," she did penalize students for failing to make corrections to assignments. This can be interpreted as disobedience or lack of diligence. Karen gave an example of a boy that accepted his poor grade, but wouldn't change it by making improvements; "he could have had an eighty-five if he made the changes I told him to make. But instead he got a seventy-five because he refused."

Karen continued, "Does bad behavior affect how I grade a kid? No. Maybe it affects how I think of him, but it won't affect how I grade him." Asked, "Do you balance behavior into grading at all?" she replied, "Not really. Usually it balances itself, because the person who is rude, arrogant, a lot of these other things, usually is not listening, and doesn't get the assignment, I mean, he balances it out himself; I don't really have to do any of that balancing."

Although Jeanette Jones explained, "Well, I don't give a grade for behavior in my classes. I guess I stopped doing that years ago," she admitted that the kind of behavior she observed on the part of a student can influence her grading. "Well. Yes, it can. You

have to be honest (laughs). I like to think I'm totally objective, but there are times when they've crossed the line, and you're not going to bend, either." She gave an example: Jeanette will let a rude, uncooperative child with a sixty-three fail, but will pass a pleasant child whose earned score is the same sixty-three.

Modeling the Moral at Mellmax High

One aspect of moral training is the modeling of hoped-for behavior. The teachers at Mellmax High strove to set a good example. In this one instance, one might say that it was the teachers' turn to be graded by their pupils. The hope that their good example would be internalized and expressed in the lives of their students was voiced by more than one of the teachers. While this modeling would have no direct effect on grades, it suggests that if the teachers are truly serious about modeling, their grading strategies would be of high moral quality, and this would be evident to their pupils.

The teachers were asked, "How do you see your role as an English teacher in the overall development of your students?" Hester Hypoint chortled, "Ha! Well…sometimes I feel like I'm trying to be a miracle worker. Um…I think what I do is very important. How successful it is, I'm not always sure, although there are students where you do definitely see progress. Others, they're hard to see that, but…..I certainly try to hit all the deficiencies…you know, work on all these areas." Modeling true diligence, Hester never gave up on a pupil, regardless of his or her grades.

Asked the same question, Gail Goodwin replied, "I had three of my very best students ask me to attend their confirmation two weekends ago, and I didn't realize they were the only kids getting confirmed in the Presbyterian Church. Well, they invited me and they wrote a beautiful message to me, and they said that in their confirmation classes they did

a lot of talk about family values and things, and 'things that we talked about in your class came up a lot, and when we talked about mentors, we all three of us agreed that you would be a good mentor'. . . And that's what it's really all about, and at least I know somebody's getting it. But it's hard too, because these are the good kids. I want to get the other ones, and maybe I am, but they just don't let me know."

Asked if this character development effort was part of the grades, Gail said, "No, that's separate from the grades. That's not something I grade. You just hope to role model, and…and sometimes it comes out in a discussion or something, where maybe they'll say something that you're like, 'Oh, something's quickened,' but that's not something you put a number on. But it can appear in essays. When they start interpreting a book…if they internalize it, and it comes out that way, then, yes, it could affect their grades."

Ellen Enders mused, "I think it [being an English teacher] is a huge role in their development, I mean, as students, as people. I'm teaching them all sorts of life lessons, and I conference with them and we spend a lot of time talking about things. You get to know the kids so well…"

Betty Browne saw herself modeling several roles. "I don't see myself as just an English teacher. Yeah, I'm going to teach them writing, and grammar: I mean, I think I'm a role model, and a coach, and a friend. …I hope I'm giving them a foundation. That's my job."

Fiona Fisher was more direct: "I like to think that I'm a positive influence on them."

Does Bad Character Mean Bad Grades?

Given the importance the English teachers at Mellmax High placed on moral training of their students, and the fact that they all claimed that English class was an excellent vehicle for moral instruction, one might expect that the display of good character in English class would be rewarded with better grades, while the display of bad character would be punished with poorer grades. But this was not so; as important as moral development was, it was kept, as far as consciously possible, out of grading by most of the teachers most of the time. Yet there were circumstances where the teachers hesitated in asserting that character never made its way into the grades.

For a "handful of kids" Luann Lender described as "hating" her, and of whom she admitted she, in turn "was not particularly fond," character did creep into the grades, but "only in the sense that, I wouldn't, after a while, put that extra effort into trying to get them to do a re-write, or encourage them to make up something. I would *tell* them, but…probably in a subtle sense, it did affect the grading. But would I penalize them and say, 'Well, I don't like you, so I'm going to give you a C on this essay?' I hope I had enough integrity not to do that." Asked if this raised an issue of fairness, she replied, "No, I don't think so. I'd tell them, but to *urge* them to do this work again would be fruitless; they weren't going to listen, so it would be like talking to a stone wall to try to encourage them. I'm not going to get a result from it, so…"

Fiona Fisher argued that she did grade the students' character, but in a different place than the report card—through the feedback she gave on their assignments. "The remarks and comments on the papers are part of the grades, it isn't just the numbers. When I write little notes on their papers, that's part of their grade. Do grades reflect a judgment

on them as people? If they received a fifty, that's going to make me believe that they're

not doing what they're supposed to be doing. I think that sometimes by writing a note on

their paper, you're letting them know that you can see what they're saying."

Fiona continued, "I'd like to say, no [character doesn't affect grades], but in the back

of your mind you're thinking, 'he was a jerk while they were working on this in class,' so

maybe it does affect the way I grade them.

"Presently (at the beginning of the year), I've just been going by the book. 'Yes,

you've been a pain in my butt, but you're doing what you're supposed to be doing, so…'

But when report card time comes, because I told them at the beginning of the year that

their attitude *does* go into their grade, then he may see a difference."

APPLYING THE MORAL FRAMEWORK TO CHARACTER DEVELOPMENT AT MELLMAX HIGH SCHOOL

The moral framework used in examining the issues of grading raised in this paper

begins with a consideration of two substructural elements in the moral life of schools that

are common entries on the list of virtues typical of character education. There is no irony

in this; truth and trust are essential both to the functioning of schools and to the

development of character. Attentiveness, too, can be seen as a kind of teacherly virtue.

Personal and individual worth are taken for granted in moral training, especially in the

caring approach, and, if courses of study were people, worthwhileness would certainly be

a virtue.

Truth and Certainty in Moral Training

We begin the application of the theoretical framework with truth and its relation to

moral training and grades. Clearly, one of the chief difficulties in moral training in

American public schools lies in the problem of truth: who can be sure what is the truth? A reflection on this problem leads one to the realization that truth in its pure form (as a virtue, as it were) depends on *certainty,* and this is what complicates moral training in schools.

Moral training must assume truth of some kind, but it cannot assume certainty without adhering to some sort of doctrine or belief system. This is not possible in a pluralistic society such as ours. Each person must "find his truth" and form his character accordingly, and no person can impose his idea of the truth on another. This is not to suggest a kind of relativism, but to recognize the constraints of our pluralistic, multicultural public school system. The philosophical argument can be complex, but belongs in another paper; for our purposes, the *assumption of truthfulness* (Jackson et al, 1993, p. 18) can guide moral education, but the inability to obtain certainty—or even to claim it—in a public school setting with regard to morality means that teachers can indeed discuss moral issues and examine and debate them, on the grounds that they are honestly doing so, but they cannot go beyond this discussion to the final step of declaring, "*This* is the moral thing to do in this situation." That decision must be left to each student to make according to his or her reason and in keeping with his or her belief system or religious doctrine, if he or she has one. One might grade the process used to get there, but not the decision itself.

Character formation is a shared process, partly the work of the teacher and partly the work of the student. It is a kind of learning, but has the added element that the learner must decide for himself what he is to be. And that means that it cannot be treated like other learning, where "the answer" can be determined with far greater certainty, and

generally agreed upon by all concerned. If this is so, then there is no grading the outcome, that is, the character of the child. One may judge it, to be sure, but this is a moral process like the one described above, and hence the judgment is a personal one. One might grade the concrete expressions of the resulting character ("shows diligence, respect, laziness", etc.), but not the character itself.

This is what we have seen in the grading practices of the teachers at Mellmax High. They hesitate to take upon themselves the right to judge the character of their students, although they do at times judge its expression, and undoubtedly privately judge their students' characters. But they do not grade them, because they cannot be "graded," since this requires some level of certainty and some sort of agreed-upon measure. They approach moral issues enthusiastically, but cautiously, and they model what they believe is good character in the hopes that it will be freely emulated.

Trust and Moral Development

Trust is an important element in any kind of instruction. Students must trust their teachers and vice versa. For each of the three kinds of moral training discussed in this chapter, trust is a prerequisite and necessary condition, and it takes slightly different forms in each of the three approaches.

In the most traditional kind of approach, the character education approach, trust on the part of the pupil takes the form of acceptance of the received list of virtues to be emulated. These must be good traits, and worthy of adoption and incorporation into one's own character. Suspicion about their value—lack of trust—can lead to either of two outcomes. First, it can lead to rejection of the virtues and failure to incorporate them. Or, perhaps worse, it can lead to the adoption of an imitation of virtue wherein the traits

are either adopted only as appearances, or they are adopted in a way which distorts them and makes them caricatures, so that fortitude becomes stubbornness, diligence becomes obsession, respect becomes flattery, and so on.

On the part of the teacher, trust demands that the pupils' display of the virtues be considered genuine, or that the teacher be so astute as to accurately identify the counterfeit versions of the virtues so that her evident skepticism is not a violation of the students' trust. This is such a difficult task that, once again, the teachers at Mellmax High shrank from attempting to grade the apparent acquisition of virtue directly. Rather, they chose to comment on it in a way that would easily show that they were judging from appearances, and could judge no deeper with certainty. Hence, the comments on the report cards and on students' papers were ways of suggesting that a low achievement grade might be explained by a lack of diligence or perseverance, an uncooperative disposition or a failure to pay attention—but that these explanations were based on appearances, and would only be correct if the teacher were astute and the student were presenting himself as he truly was.

For the second approach to moral training, the moral education model, trust demands that both the teachers and the students explore moral issues frankly and as fully as they are able, without resorting to sophistry in their arguments for the morality of one course of action or another. The students have to be able to trust the teacher to provide a full consideration of the matter, and not to withhold some part that would alter their decision about the right course to follow. In addition, the students have to be able to trust the teacher to provide what she believes is, as Luann Lender claims, "a way of looking at the world and at life." This kind of trust may allow for grading of the *application* of this

"way of looking at the world," but it does not make it possible to grade the decisions that result; if this were so, what has been presented is not "a way" of looking at the world, but "the teacher's way."

In the third model, the caring approach to moral development, trust assumes an even greater importance. Because the caring approach is relational, it depends on a bond of trust between the teacher and the students, and among the students and the school community. This kind of trust makes caring both possible and genuine, and means that the person is valued more highly than his achievements, his virtues, or his character. It may make any kind of grading that judges persons—or their character—a violation of that trust. On the other hand, it may *require* such judgment, since the bond of trust among all the members of the school community requires the development of suitable character in those members.

Here we see again that moral development can at most only give color to grading practices, so to speak. Since moral development cannot be judged directly or with certainty, the English teachers at Mellmax High School, while perhaps letting their perception of a child's character push a grade slightly up or down, as a rule resisted the assessment of character development in a quantitative way.

Worthwhileness and Character Development

The worthwhileness of character development is evident in both its long tradition and the obvious enthusiasm with which the English teachers at Mellmax High embraced the moral discussion made possible through literature. That the development of character is important enough to merit considerable attention in schooling is a premise of every

214

author who has championed one school of thought or another regarding the best way in which to go about establishing character in our youth.

The one question that arises out of a consideration of the worthwhileness of character training and grading is an ironic one: if character development is so worthwhile, perhaps even the most important purpose of the American public school, why isn't it important enough to be graded? Americans grade everything they think of as important or worthwhile, except character. The recognition of the worthwhileness of character development, whether in terms of character education, moral training, or the caring approach, makes it a tempting target for grading.

Of course, we have already seen reasons why character cannot be graded; we have no sure method, standard, or device for measuring it. At most, the teachers at Mellmax High use judgments about character development in limited and indirect ways when grading their students. Yet we all recognize good character, are repulsed by bad character, and wish our children to develop true character. Perhaps it is a tribute to the depth of its worthwhileness that character development pervades the moral life of schools and yet remains elusive and immeasurable.

Intellectual Attentiveness—Juggling Multiple Inputs

The exploration of moral issues requires close intellectual attentiveness on the part of the teacher involved; the development of students' moral reasoning, whether in keeping with Kohlberg's model or otherwise, requires careful attention. Especially since the usual vehicle for such exploration is classroom discussion, the teacher may need to be able to follow several lines of reasoning simultaneously as different students contribute in different ways to the advancement of the discussion. If the skills the teacher purports to

impart are designed to allow near-universal applicability for inquiry into moral issues and dilemmas, then these must be modeled effectively in the moment, an exercise in attentiveness that could be very taxing.

In addition, if one situation can be better understood by comparison with other situations, perhaps real or perhaps hypothetical, the teacher must be able to keep all of this in play for at least the duration of the discussion, and possibly for several days— perhaps for the duration of a complete unit of instruction. As Alice Andrews pointed out, discussion of moral issues can involve drawing connections with many sources—other novels, personal situations, and other ideas—and this requires intellectual attentiveness to many different ideas all at once.

Add to this the fact that a class may consist of twenty-five or more students, and the burden of intellectual attentiveness becomes quite intimidating. While it is unlikely that all twenty-five or more students will participate fully at any given moment, the number of possible lines of reasoning—both sound and unsound—that may be presented could be staggering. Perhaps here, more than anyplace else in the course of English instruction, the demands of intellectual attentiveness become obvious and taxing.

Here we are talking about one approach to moral training; in the character education approach, the need for intellectual attentiveness is slightly different, but no less demanding; being fully familiar with the virtues and their variations, their false cousins and imitations, and keeping all this in order while helping the students become familiar with the virtues and their expressions in action is another daunting task, but one which the teacher must shoulder if character education is to be successful.

In the caring approach, where personal relationships are stressed, the demands of intellectual awareness require that the teacher be truly familiar with the intellectual features and idiosyncrasies of his students so that he can both nurture the caring relationship and foster the intellectual advancement of each of his pupils.

In the matter of grading, some teachers will assign points for participation in class discussions. This may amount to little more than a device to encourage students to contribute, and may not even actually be kept track of, as long as the discussion is lively. Or it may mean taxing one's intellectual awareness even further in order to keep track of who actually participates, at what level, and whether this is in keeping with perceived ability and expectations for the individual student.

Another area for grading, as pointed out by Gail Goodwin, is the essay that may apply the technique for examination of a moral dilemma modeled in moral education, present a personal moral stance, or evaluate the morality of some action in a novel or real life. Here the teacher's intellectual attentiveness must be focused on the individual student, and the teacher must carefully maintain the neutrality she advocates during modeling in order to accurately judge the process the student is following in arriving at his or her conclusion. As Mary Minton pointed out, "You have to be careful, because they may have very different ideas from your own."

Moral Attentiveness—Paying Heed to Development

Moral attentiveness calls for the teacher to pay heed to the persons his or her students are becoming, and so the area of moral development is a particularly taxing one when considering the demands of the classroom and the relations between teacher and students. If a teacher is truly morally attentive, all that was true for the above discussion on

intellectual attentiveness applies again in a new light; whereas intellectual attentiveness may require keeping track of multiple cases of mental development all at once, moral attentiveness requires careful attention to multiple examples of moral conduct all at once and all the time. Every student is a moral actor at every moment, and the teacher must be attuned to the students' actions and words throughout the duration of classes, and at any time when he or she may be in the position to observe the students—in the halls, in the gym or cafeteria, and even out in the community.

In terms of grading, one can see why teachers would shy away from drawing upon moral attentiveness to give a grade even as one can see how it might inform some part of the grading process. On the one hand, the opportunity for actually observing moral conduct is constant, but the likelihood of any given observation's representativeness of the moral character of the student is very small. On the other hand, these occasional observations can make deep impressions, and keeping those entirely out of the grading picture can be very difficult, no matter how impartial or neutral the teacher believes himself to be. While classroom grades may be only slightly affected by such things, decisions about placement in service organizations or the national honor society, for example, can be based on a single observation of good or bad behavior by a single teacher in a single moment. Moral attentiveness, then, can be seen as both an essential trait for a good teacher and a terrible moral burden.

English teachers, whose course content provides numerous useful vehicles for the examination of moral issues, discussion of right and wrong, and the development of reflective thinking on such ideas, traditionally view the formation of their students as part and parcel of their duties, but the idea of grading such formation makes most teachers

uneasy. For some of the English teachers at Mellmax High, the moral development of their students had some subtle influence on their grading practices. For most, the many complications of letting such things "creep in" to one's grading system prohibited them from factoring the moral development of their students into their grades in any conscious way. While the character development of students undoubtedly affects their behavior and the way they approach their studies, and thus affects their grades, it does so indirectly. The English teachers of Mellmax High prefer to determine achievement grades apart from inferences about students' moral development, dealing with judgments about character development through narrative grades, commentary, and personal interaction with their students.

CHAPTER SIX: CONCLUSION

Grading practices in American high schools have been a topic of perennial interest (Laska & Jaurez, 1992; Brookhart, 2004). Much of what goes on in the classroom to inform the gradebook seems baffling to the measurement experts, who decry what they see as an ill-chosen hodgepodge of bits and pieces, apparently collected willy-nilly and without careful thought. Yet this study demonstrates that this is not so: the English teachers who participated gave considerable thought to the systems they employed for arriving at grades, the component pieces they used as graded assignments, and the ways in which these choices played out as the school year progressed. The use of the phrase "hodgepodge grading," first used by Brookhart (1991) and subsequently taken up by Cross and Frary (1999) and many others, is unfair; it devalues the diligent efforts of English teachers everywhere and denigrates their professional wisdom and judgment.

CENTRAL QUESTIONS

The central questions with which this study began have been explored, and answers can be offered based on the information collected from the English department at Mellmax High School. These are not, however, definitive and final answers, but initial answers to questions that can be answered correctly in many different ways. One purpose of this study was to reveal the complexity, rather than the apparent chaos, of grading in high school English. Another was to spark interest in the moral dimensions of grading, chiefly in English, but in other subjects as well.

How do English Teachers Decide What is Right in Giving Grades to Individual Students?

We began by asking, "How do English teachers decide what is right in giving grades to individual students?" The question has two meanings: "How do the teachers decide what is the correct way to arrive at grades?" and "How do teachers decide what is the morally correct way to arrive at grades?"

Deciding How to Arrive at Grades

This study answers the first part of the initial question by revealing that teachers decide on the correct way to arrive at grades through structure, purpose, and an understanding of the constraints imposed on them as part of a larger system. This combination of sources informs the teacher's creation of a grading system which incorporates as many of the elements as the teacher believes are important and manageable; hence, some teachers "grade everything" (including non-achievement factors), others "grade some things" (selected assignments and non-achievement factors), and others "grade only what is assigned."

Structure has to do with the arrangement of the curriculum, the classroom, and received culture of schooling. Part of this is within the control of the teacher and part is not.

The curriculum is partly based on what is received from, and required by, the school or the department, and partly up to the English teacher to decide. This plays into grading because the level of difficulty and content affect grading. More importantly, the way the

teacher tests and the kinds of assignments given to aid students in mastering the content have a great deal to do with what goes into the grading.

The classroom and its rules and routines are incorporated into a teacher's grading system in many ways. Some of these can be obvious, as when a teacher employs a system of rewards and punishments for behavior in class that either directly affect grades by giving or taking away points, or indirectly affects them by giving a "pass" so some assignment need not be completed and hence is not calculated into the average. In more subtle ways, as we have seen in this study, observation of class behavior may lead teachers into a practice of "point-shaving," taking little deductions on assignments from the uncooperative and giving little bonuses to the exemplary student. Some teachers feel that behavior that interferes with instruction is rightly penalized via the grading, since it impedes the progress of the class as a whole, and usually that of the individual responsible as well.

Purpose in grading is not single, but multiple; teachers grade with many underlying purposes in mind (Stiggins & Conklin, 1992; Karmel, 1970; Brookhart, 2004). This study has confirmed many of these purposes: assessment and evaluation; establishment of expectations; sizing up the pupils; making judgments about appropriate next steps in instruction; motivation and encouragement; practice; even development of character. These are simultaneously individual and collective, i.e., aimed at specific students and at whole classes.

Focusing single-mindedly on one purpose and neglecting others may streamline grading, but gives an incomplete picture; teachers are cognizant of this fact, and so attempt to do many things at many levels. This impacts their grading practice and

requires that, regardless of its apparent stability of structure, it must be adjusted and modified continuously. One might think of a grading system as a kind of organic entity or living thing, constantly changing and yet always the same in some deeper sense.

An understanding of the constraints that result from being part of the larger school system structure means that teachers must take the school district's grading system into account in their own system. The grading systems explored in chapter three highlight the ways the larger system can influence grading. Perhaps the most obvious of these is the "grading cycle," which dictates when grades are to be determined. This in itself affects the grading outcomes, since it fixes grades at set moments, at times favorable to a given student, at other times unfavorable. Teachers, aware of this dilemma, must both submit to the cycle and find ways to counter its negative consequences.

Another constraint is as obvious as the form the grades must take and the limits imposed by the report form: at Mellmax High, a single number grade must capture most of the information the teacher has amassed for her students, and three pre-scripted comments must suffice for remarks or explanations. Knowing that this is all that the parents receive, the teachers at Mellmax High often used the comments to help explain the number grade. Rather than using the comments to present non-achievement factors and the number grade to average only achievement scores, the number grade served as a composite of all factors in a teacher's grading system, and the comments shed light on what might have skewed the outcome, whether it was low test scores (achievement) or talking back (non-achievement).

Deciding How to Give the "Right Grade"

In regard to the second meaning of the first question, the moral meaning, each teacher determines the morally correct way to arrive at grades through a combination of philosophy, understanding, and personal beliefs, and a process of reflection on all these things. Depending on the individual, this process of reflection might be deep or shallow; it might occur only when grades are due, or only during times away from school; it might be occasional, and prompted by an immediate grading issue, or it might be a regular process that is nearly continuous. The interviews were timed so that they took place at the times when grades were due, on the premise that this would be the time when the teachers would be most likely to be reflecting on the way they grade and whether the grades they ended up giving were in fact in keeping with what they felt were the "right grades."

Like most issues of moral weight, some level of doubt in the final decisions was always present, yet decisions had to be made, and were. But these were not made lightly, and were informed not merely by the concrete data of the graded papers, but also by the individual teacher's philosophy of teaching and learning, an understanding of the learners—despite its imperfect and limited nature, the teacher's personal beliefs about what is good and good for the pupils, and an attitude of caring—caring for the pupils, for the school, and for the community.

How and What do English Teachers Intend to Communicate to Students Through Grades?

The English teachers at Mellmax High used their grades to communicate many things, in keeping with the multiple purposes listed above that guided their teaching.

While there is a tendency for the public to think of grades in the very limited sense of the report card grades issued at regular intervals, teachers are giving grades on many different kinds of assignments and making judgments through observation constantly (Airasian & Jones, 1993; Wiggins, 1994). The requirement that all this information be converted into a single number and a handful of comments at the end of each marking period, however, means that the teachers are forced to consider the report card grade as a rough composite of all they have collected over the period preceding the report.

This study revealed that the English teachers at Mellmax High were very conscious of the limitations imposed by the grade report system of their school, and went to considerable lengths to conference with their students and explain their grades, both collectively and individually. They explained that there were many, many factors being worked into the grade in some fashion, and were confident that, by and large, the students understood what they were attempting to communicate through their grades.

Through numbers that are interpreted through a shared understanding, narrative grading that addresses non-quantitative features of the student's work and behavior, and personal interaction with the students as individuals, the teachers communicate an overall assessment of progress in terms of achievement, growth, and personal accomplishment unique to the individual student and usually understood (according to the teachers) by the student. This overall assessment, the Mellmax High English teachers stressed, is their *own* assessment of the student, and again, the students, by and large, understood that another teacher might see them differently. Despite the official nature of the school district grade report, the personal nature of grades and grading is part of the shared understanding of grades that governs high school grading practices.

Is Character Development a Part of the Role of an English Teacher, and if so, does this Influence Grading, and How?

Clearly, all of the English teachers at Mellmax High believed that their students' character formation was part of their role. This is in keeping with a long tradition of American public schools, wherein literature is to serve as a character-building force (Kliebard, 1995; Gilness, 2003; Edgington, 2002). The study of literature invites discussion of the moral, and this is part of character development. While the teachers at Mellmax used an eclectic approach to character development, all three of the major strands could be detected in their practice.

Character education with its list of virtues (Bennett, 1993; Kilpatrick; 1992) was evident especially in remarks about classwork and behavior that could be directly observed, like task-orientation or teamwork. Moral education (Kohlberg, 1981, 1984; Dewey, 1909, Sharp, 1917) was noticeable in teachers who argued that they were giving the students "a way of looking at life." And the caring approach (Noddings, 1992, 2002; Sizer & Sizer, 1999) was especially evident in the remarks of the junior high teachers, who were clearly concerned with both the grades their students were getting and the ways in which they were adjusting to the new demands of high school. None of the teachers restricted their approach to a single method, but used elements of any of the three as they felt appropriate.

Although the teachers had strong feelings about the importance of character development and their role in it, their efforts were, for the most part, kept apart from their grading. This was partly due to concerns about imposing personal views on their students, and partly because of the difficulty in judging character with any certainty. How one behaves in school may or may not be consistent with how one behaves outside

of school; in addition, the teachers viewed their pupils as "works in progress," so that development varies widely from student to student.

Nonetheless, because certain assignments, especially homework and work done in class, could be interpreted in terms of their demonstration of some virtue (diligence, perseverance, and so on) or desirable trait (a strong work ethic; courtesy and consideration for others, etc.) which affected its grading, the argument can be made, and is here, that character development issues do creep into grading in subtle ways, sometimes in ways that the teacher is aware of, but more often in ways the teacher does not recognize as reflecting an assessment of a student's character.

ADDITIONAL FINDINGS

In addition to shedding light on the grading processes of the English teachers at Mellmax High, this study provided insight into the effects of school district grading requirements on the teacher's methods of grading, into teacher expectancies, and into the teachers' understanding of their role in the character development of their students.

Influence of School District Grading Requirements

School district grading requirements influence teachers' grading practices in several ways. In addition to the effect of the school calendar and the grade report cycle, implicit understandings about acceptable rates of failure influence grading, especially for the untenured teachers and those teaching seniors, whose graduation would depend on passing 12th grade English. While there were no explicit rules regarding the number of failures, all the teachers mentioned pressures from the office regarding these numbers, and all had some level that they had inferred was the point beyond which additional failures would not be countenanced by the main office.

Policy regarding minimum grade levels for the first semester (no grade lower than a fifty) were again implicit, evidently a result of long-standing tradition in the district. This received "rule" was widely flouted by the English department because it ran counter to a sense of justice. However, some teachers followed the rule out of a sense of compassion for those students they felt were trying.

In addition, it became clear that district grading practices shaped the way grades were interpreted by the community, and that the realization that this was so meant that the widely understood meaning of grades had to be factored into the way teachers graded. ("If I give this kid a ninety and this one an eighty-five, everyone will think that I'm being harsh" or "If I give this kid a passing grade, but not this one, everyone will think I'm being unfair to athletes.") Thus the community itself exerted subtle pressures on the teachers with regard to grading.

When an individual teacher wished grades to be understood in a way different from that generally shared within the community, the teacher spent some time explaining his or her grading scheme and its rationale to his students—and sometimes to their parents. In general, grades carry greater meaning to the teachers than to the students, greater meaning to the students than to their parents, and greater meaning to parents of children in school than to the wider community.

Teacher Expectations

This study confirmed part of the research in teacher expectancies (Dusek, 1985), but did not support other parts. It confirmed that teachers form expectations and that these influence grading. However, it did not support the contention that teachers are capricious

in forming expectations, or can be easily influenced by unsupported information (Rosenthal & Jacobson, 1968) about a student's potential for success or failure.

Indeed, the evidence suggests that most teachers are scrupulous about maintaining accurate expectations and require new information on a near constant basis to do so. This information comes from many different sources and the teachers at Mellmax High adjusted their expectations in keeping with their analysis of the information they collected.

In addition, this study indicates that teachers work with their students to arrive at shared understandings about attitude and effort. These are usually communicated early in the year, and reinforced regularly through whole-class reminders and individual conferences, especially with students who are not living up to expectations.

Character formation

While the English teachers at Mellmax High School believe that the character development of their students is part of their duty and work consciously at it, they do not believe it should be graded because of its highly personal and sensitive nature. This is partly because the grading structures of schools are not designed for it, and partly because of the pluralistic nature of our society, which discourages comparisons among students of a relatively direct kind in matters of morals or character.

Although all of the teachers felt that character formation was extremely important in high school, not one believed it should be graded in itself, even if it were to be awarded its own category apart from the rest of the students' grades. In addition, none of the teachers indicated that it should form the content of a separate course, despite the legislated requirement for such a course in New York State. In fact, they felt that English

class was a better forum than a separate course would be because it provided a context for the moral issues discussed and made relating moral issues to other areas of the students' lives easier.

LIMITATIONS

While the entire English department of Mellmax High School participated in this study, and, as noted in the second chapter, the school shares many similarities with other upstate New York districts, every school is unique. As such, although the curriculum and general profile of the Mellmax district is a typical one, it is not representative in the scientific sense. As is true in any qualitative research drawing upon a purposive sample, one may safely say that insofar as the Mellmax district and its teachers are similar to another district and its teachers, what is true for one may be true for the other.

The purposive nature of the sample means that it is one that makes sense for comparison with others like it, but what is true of one may or may not be true of the other. It may be very similar, only slightly similar, or not at all similar. The value of the study derives chiefly from the insight it offers into the grading practices of the teachers involved, not the extent to which any particular part of the study may be transferred to another district or group of teachers on the basis of similarity.

A second limitation results from the almost exclusive use of interviews as the method of data collection. Here there are two observations to be made: first, interviews allow the informants to present themselves as they wish to be presented. While one has no reason to doubt the veracity of the teachers' responses, it is natural for people to present themselves in the best possible light, so that some practices may be more or less extensive than depicted, and there may actually be times when teachers do not follow

their own rules as scrupulously as they would like. Secondly, there is also the possibility that the teachers honestly believe they do exactly what they say they do, and yet act in ways that are in fact slightly different. The absence of observational data to support the self-reported practices of the teachers somewhat weakens the study.

Still another limitation may derive from the nature of qualitative studies like this one that amass large volumes of data and require careful sorting in carrying out a comparative analysis. In the end, the data presented here amounts to only a small part of the total collected. While it is reported accurately and the interpretations are defensible and reasonable, one could argue that other elements from the data could have been presented and that those could create a somewhat different picture.

Much like a lawyer building a case, a qualitative researcher must decide which pieces of "evidence" are most important, which are relevant and which are not, and assemble them in a way that will be coherent and compelling. However, a lawyer is in fact attempting to persuade belief in favor of his client; the researcher aims at presenting an accurate and comprehensive picture of the case being studied without "slanting" the data by leaving out relevant pieces or ignoring hard-to-explain data elements. Qualitative analysis requires the adoption of a perspective which affects the way data is understood; while the researcher must avoid bias in sorting and reporting the data, he cannot escape from his own way of looking at things altogether, and some subtle biases almost certainly color his approach and influence his choices of what to include and what to leave out. Every effort has been made to avoid this as far as possible in this study, but the limitation remains.

The vast amount of data mentioned also suggests another limitation: the picture presented may not be as complete or detailed as possible. This limitation, however, can be corrected through further research.

IMPLICATIONS

Research on the topic of grading has in recent years taken an increasingly narrow view of assessment, adopting a perspective which limits the scope of measurement to those things easily tested with paper and pencil tests. This has had negative effects on the general understanding of the role assessment plays in schools because it has encouraged the neglect of many developmental issues. It has also, perhaps, helped to feed the testing frenzy currently characterizing school reform movements. The transformation of high stakes tests from one source of data among many into the only one worth having is a distortion of their value.

The measurement community should not adopt a "tunnel vision" approach to grading, simply ignoring those areas that teachers, parents, and communities believe are important and should somehow be incorporated into the grades a child receives. The holistic nature of the report card grade is widely accepted; grades are understood and accepted as a composite of many, many things, and attempting to narrow them to a limited set of indicators that reflect only a small number of sources of information, and only very limited kinds of information from those sources, is a disservice to all parties involved, including the measurement community itself.

Teacher education has long been found wanting in preparing teachers to grade their students; while assessment and evaluation courses have become a common requirement, these tend to be driven by the perspective of the measurement community and suffer from

a lack of appreciation of the constraints of the school system and the realities of the classroom. A more thorough understanding of the total picture of grading in schools would be beneficial to aspiring teachers.

The actual grading process is deeply moral and colored by moral considerations that make assessment which is a reduction to pure "achievement" measures alone an almost unethical practice. The "reflective practice" movement of a decade ago revealed the need to stimulate and develop reflection in new teachers, and some of that reflection should be directed toward the moral issues involved in grading. Grading systems devoid of moral issues are suitable only for quality inspections of washing machines and the like; people are moral beings, and must be judged accordingly. Certainly the moral issues of grading add complications, but that is part and parcel of being human and enjoying the capacity for moral action.

FUTURE RESEARCH

Several areas of grading involving moral issues have been exposed in this study and merit further research. First, a larger scale inquiry into school district grading demands and the way these shape the grading practices of the teachers would be well worthwhile. With states like New York adopting "school district report cards" issued by the state education department and modeled on school grade report forms, understanding how the requirements of data reporting can shape the data itself—by narrowing the indicators based on ease of collection, by limiting the kinds of data to those things easily expressed in numbers, and so on—would be of great value. The moral issues of judging by report card need to be more thoroughly explored if this is to become a more and more widely adopted practice.

While there have been many studies of the kinds of things teachers grade and the reasons they think they should be graded, most of these studies have lacked the dimension of the moral issues involved. Instead, the various sorts of data teachers collect are judged by utilitarian standards—this is easily assessed, so use it; that is hard to judge or quantify, so abandon it, and so on. This reduces schooling to a process of mastering skills and collecting "knowledge," and neglects the human development issues which are arguably far more important to an amicable society and effective citizenship.

At the same time, there is considerable research into the area of schools and a just society, character education, peace education, and so on. Little of this, however, explores grading issues in any depth. More research that keeps the two together is called for, since schools are in fact attempting to pursue both areas at once—indeed, they are expected to.

Future research on grading should be cognizant of the moral issues deeply embedded in schooling and grading; these issues are so pervasive that ignoring them or studying grading apart from them gives a false understanding, and contributes to the mistaken interpretation of high school teachers' grading practices as capricious and confused. In addition, it narrows the scope of instruction and encourages the neglect of areas of great importance which are not easily measured—areas like moral development.

The aim of future research should be a fuller understanding of the grading process rather than an attempt to impose an idealized, "scientifically sound," system on an already established way of doing that has many sound elements already in place. Grading and grading practices need to be studied for what they are, not according to a preconceived notion of what they ought to be. Suggestions for changing grading methods should be rooted in the realities of the classroom and the politics and social frameworks

of the schools. Grading cannot be fully understood apart from its complete context. Future research should aim at gaining a broader understanding of that context and its moral dimensions in order to improve the practices that govern the way grades are determined.

WORKS CITED

Adams, W. L. (1932). Why teachers say they fail pupils. *Education, Administration, and Supervision, 18,* 594-600.

Aiken, L. R. (1963). The grading behavior of college faculty. *Educational and Psychological Measurement, 23,* 319-322.

Airasian, P. (1996). *Assessment in the classroom*. NY: McGraw Hill.

Airasian, P. W. & Jones, A. M. (1993). *The teacher as applied measurer: Realities of classroom measurement and assessment.* Applied Measurement in Education, 6, 241-254.

Anderson, L.W., & Bourke, S. F. (2000). *Assessing affective characteristics in schools*. Second edition. Mahwah, NJ: Lawrence Erlbaum.

Baier, A. C. (1995) *Moral Prejudices: Essays on ethics.* Cambridge, MA: Harvard University.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.

Barber, T. X. (1969). Five attempts to replicate the experimenter bias effect. *Journal of consulting and Clinical Psychology, 33 (1),* 1-14.

Bass, B. M. (1951). Intrauniversity variation in grading practices. *Journal of Educational Psychology, 42,* 366-368.

Bells, W. C. (1930). Reliability of repeated grading of essay type examinations. *Journal of Educational Psychology, 21,* 48-52.

Bennett, W. J. (Ed.) (1993). *The book of virtues*. NY: Simon & Schuster.

Berkowitz, M. W. (2002). The science of character education. In W. Damon (Ed.), *Bringing in a new era in character education* (pp.43-63). Stanford, CA: Hoover Institution Press.

Beyer, L. (1997). The moral contours of teacher education. *Journal of Teacher Education, 48(4),* 245-254.

Bloom, B., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on summative and formative evaluation of student learning.* NY: McGraw-Hill.

Bogdan, R. C. & Biklin, S. K. (1998 [1982]). *Qualitative Research for Education: An introduction to Theory and Methods,* 3rd ed. Boston: Allyn & Bacon.

Borko, H. & Shavelson, R. J. (1978). Teachers' sensitivity to the reliability of information in making causal attributions in an achievement situation. *Journal of Educational Psychology, 70,* 271-279.

Bracey, G. W. (1994). Grade Inflation? *Phi Delta Kappan, 76 (4):* 328-329.

Briggs, C. L. (1986). Learning to ask: *A sociolinguistic appraisal of the role of the interview in social science research.* New York: Cambridge University Press.

Brookhart, S. M. (1991). Grading practices and validity. *Educational Measurement: Issues and Practice, 10(1),* 35-36.

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7,* 279-301.

Brookhart, S. M. (2004). *Grading.* Upper Saddle River, NJ: Pearson Education.

Brophy, J. & Good, T. (1969). *Teacher-child dyadic interactions: a manual for coding classroom behavior.* Report Services no. 27 Austin, TX: University of Texas at Austin, Research and Development Center for Teacher Education.

Brophy, J. C. and Good, T. L. (1970). Teachers' communications of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology, 60,* 365-374.

Buckley, J. (1968). Who is Pygmalion, which is Galatea? *Phi Delta Kappan, October*, p.124.

Buzzelli, C. A. & Johnston, B. (2002). *The moral dimensions of teaching: Language, power and culture in classroom interaction*. New York: Routledge/Falmer Press.

Clinton, H. R. (1996). *It takes a village.* New York: Simon and Schuster.

Cooper, H. (1977). Controlling personal rewards: Professional teachers' use of differential feedback and the effects of feedback on the student's motivation to perform. *Journal of Educational Psychology, 69 (4),* 419-427.

Cooper, H. (1979). Pygmalion grows up: A model for teacher expectation communication and performance influence. *Review of Educational Research, 49*, 389-410.

Cooper, H. M. (1985). Models of teacher expectation communication. In Dusek, J., ed., *Teacher Expectancies*. Hillsdale, NJ: Lawrence Erlbaum, pp. 135-158.

Cooper, H. M., & Baron, R. M., (1977). Academic expectations and attributed responsibility as predictors of professional teachers' reinforcement behavior. *Journal of Educational Psychology, 69,* 409-418.

Cooper, H. M. & Burger, J. M. (1980). How teachers explain students' academic performance: A categorization of free response academic attributions. *American Educational Research Journal, 17,* 95-109.

Cooper, H., & Good, T. (1983) *Pygmalion grows up: Studies in the expectation communication process.* NY: Longman.

Covington, M. V., Spratt, M. F., & Omelich, C. L. (1980). Is effort enough, or does diligence count too? Student and teacher reactions to effort stability in failure. *Journal of Educational Psychology, 72*, 717-729.

Cross, L. H. & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education, 12(1),* 53-72.

Cureton. L. W. (1971). The history of grading practices. *Measurement in Education, 2(4),* 1-8.

Delattre, E., & Russell, W. (1993). Schooling, moral principles, and the formation of character. *Journal of Education, 175 (2),* 23-43.

Davis, M. (2003). What's wrong with character education? *American Journal of Education, 110,* 32-57.

Denzin, N. K. & Lincoln, Y. S., (Eds.) (1994). *Handbook of qualitative research.* Thousand Oaks, CA: Sage.

Dewey, T. (1909). *Moral Principles in Education.* Carbondale: Southern Illinois University Press.

Dewey, T. (1916). *Democracy and education*. NY: Free Press.

Dexter, E. S. (1935). The Effect of fatigue or boredom on teachers' marks. *Journal of Educational Research, 28,* 664-667.

Dockery, E. R. (1995). Better grading practices. *Education Digest, 60 (5):* 34-37.

Dusek, J. B., Ed., (1985). *Teacher Expectancies*. Hillsdale, NJ: Lawrence Erlbaum.

Ebel, R. L., and Frisbie, D. A. (1986*). Essentials of Educational Measurement (4th ed.).* Englewood Cliffs, NJ: Prentice-Hall.

Edgington, W. D. (2002). To promote character education, use literature for children and adolescents. The Social Studies, May/June, 113-116.

Elashoff J. D. & Snow, R. E. (1971). *Pygmalion reconsidered.* Worthington, OH: Jones.

Entwistle, D., & Webster, M. (1972). Raising children's performance expectations. *Social Science Research 1,* 147-158.

Evans, F. B. (1976). What research says about grading. In S. B. & J. A. Bellanca (Eds.) *Degrading the grading myths: A primer of alternatives to grades and marks,* pp. 30-50. Washington: Association for Supervision and Curriculum Development.

Feldman, A., Alibrandi, M., & Kropf, A. (1998). Grading with points: The determination of report card grades by high school science teachers. *School Science & Mathematics, Mar. 98, Vol. 98 (3),* 140-149.

Fraenkel , J. R. & Wallen, N. E. (2000) *How to design and evaluate research in education,* fourth edition. NY: McGraw Hill.

Gay, L. R. & Airasian, P. (2003). Educational research: Competencies for analysis and application, 7th edition. Upper Saddle River, NJ: Pearson Education.

Gecan, C., & Mulholland-Glaze, B. (1993). The teacher's place in the formation of students' character. *Journal of Education 175 (2),* 45-57.

Gephart, W. J., and Antonoplos, D. P. (1969). The effects of expectancy and other research-biasing factors. *Phi Delta Kappan, 50 (10),* 579-583.

Gibbs, L. J., & Earley, E. J. (1994). *Using children's literature to develop core values.* (Report No. ISBN-0-87367-362-X). Bloomington, IN: Phi Delta Kappa Educational Foundation. (ERIC Document Reproduction Service No. ED 366992).

Gilligan, C. (1977). *In a different voice: Women's conceptions of the self and morality.* Harvard Educational Review, 47, 428-517.

Gilness, J. (2003). How to integrate character education into the curriculum. *Phi Delta Kappan, November*, 243-245.

Giroux, H. & Purpel, D. (Eds.) (1983). *The hidden curriculum and moral education: Deception or discovery?* Berkeley, CA: McCutchan Publishing Corp.

Glaser, B. & Strauss, A. L. (1967). *The discovery of grounded theory: strategies for qualitative research.* Chicago: Aldine.

Glasser, W. (1971). *The effect of failure on the life of a child*. Washington: National Association of Elementary School Principals.

Goodlad, J., Soder, R., & Sirotnik, K. (1990). *The moral dimensions of teaching.* San Francisco: Jossey-Bass.

Grimes, L. (1981). Learned helplessness and attribution theory: Redefining children's learning problems. *Learning Disability Quarterly, 4,* 91-100.

Gronlund, N. E. (1985). *Measurement and evaluation in teaching.* NY: Macmillan.

Hall, V. C. & Merkel, S. P. (1985). Teacher expectancy effects and educational psychology. In *Teacher Expectancies*, J. B. Dusek, ed., Hillsdale, NJ: Lawrence Erlbaum, pp. 67-92.

Halleck, R. P. (1905) *Journal of Proceedings and Addresses of the Forty-ninth Annual Meeting of the National Association of Educators,* p. 869.

Hansen, D. T. (1997). *Understanding students.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, Illinois, Mar. 27, 1997.

Hargis, C. H. (1990). *Grades and grading practices: Obstacles to improving education and to helping at-risk students.* Springfield, IL: Charles C. Thomas.

Harris, W. T. (1901). Recent growth of public high schools in the United States as affecting the attendance of colleges. *Journal of the Proceedings and Addresses of the Forty-Sixth Annual Meeting of the National Education Association, 1901,* 175.

Heider, F. (1958). *The psychology of interpersonal relations.* NY: Wiley.

Hendrickson, J. & Gable, R. (1997). Collaborative assessment of students with diverse needs: equitable, accountable, and effective grading. *Preventing School Failure, 41 (4):* 159-164.

Hopkins, K. D., Stanley, J. C., & Hopkins, B. R. (1990). *Educational and psychological measurement and evaluation (7th ed.).* Englewood Cliffs, NJ: Prentice-Hall.

Horrocks, J. E., & Schoonover, T. I. (1968). *Measurement for Teachers.* Columbus, OH: Chas. E. Merrill Publishing.

Howard, R. W., Berkowitz, M. W., & Schaeffer, E. F. (2004). Politics of Character Education. *Educational Policy, 18 (1),* 188-215.

Jackson, Boostrom, & Hansen (1993). *The moral life of schools.* San Francisco: Jossey-Bass.

Jensen, A. R., (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39,* 1-123.

Jones, E. E., Rock, L., Shaver, K.G., et al, (1968). Pattern of performance and ability attribution: An unexpected primacy effect. *Journal of Personality and Social Psychology, 10,* 317-340.

Joseph, P. B., & Efron, S. (2005). Seven worlds of moral education. Phi Delta Kappan, March, 525-533.

Kadamus, J. A. (2001) Memo to district superintendents of Schools, 9/2001. Albany, NY: State Education Department/The University of the State of New York.

Karmel, L. J. (1970). *Measurement and evaluation in the schools.* Toronto: Collier-MacMillan Canada.

Kelley, H. H. (1967). Attribution theory in social psychology. *Nebraska Symposium on Motivation, 15,* 192-238.

Kelley, H. H. (1972). *Causal schemata and the attribution process.* Morristown, NJ: General Learning Press.

Kerr, D. (1987). *Authority and responsibility in public schooling.* National Society for the Study of Education.

Kilpatrick, W. (1992). *Why Johnny can't tell right from wrong: Moral illiteracy and the case for character education.* NY: Simon & Schuster.

Kirby, B. C. (1962). Three error sources in college grades. *Journal of Experimental Education, 31,* 213-218.

Kirschenbaum, H., Napier, R. W. & Simon, S. B. (1971). *Wad-ja-get? The grading game in American education.* New York, NY: Hart Publishing.

Kliebard, H. M. (1995). *The struggle for the American curriculum, 1893-1958*, second edition. NY: Routledge.

Kohlberg, L. (1981). *Essays on moral development. Vol. 1: The philosophy of moral development.* San Francisco: Harper & Row.

Kohlberg, L. (1984). *Essays on moral development. Vol. 2: The psychology of moral development.* San Francisco: Harper & Row.

Lang, J. M. I., (2004). Failing to motivate. *Chronicle of Higher Education, 51 (15),* C2-4.

Laska, J. A., & Juarez, T. (Eds.) (1992). *Grading and marking in American schools: Two centuries of debate.* Springfield, IL: Charles C. Thomas.

Lewin, K. (1935). *A dynamic theory of personality*. NY: McGraw-Hill.

Lewin, K. (1936). *Principles of topological psychology*. NY: McGraw-Hill.

Lewin, K. (1938). *The conceptual representation and the measurement of psychological forces.* Durham, NC: Duke University.

Lickona, T. (1991). *Educating for character*. NY: Bantam Books.

Lickona, T. (1998). Character education: Seven crucial issues. *Action in Teacher Education, 20 (4),* 77-84.

Martin, G. H., (1899). Unseen Force in Character-Making. *Journal of Proceedings and Addresses of the Forty-third Annual Meeting of the National Association of Educators,* pp.279-286.

Martinson, D. L. (2003). High school students and character education: It all starts at Wendy's. The Clearing House, 77 (1), 14-17.

McArthur, A. (1972) The how and what of why: Some determinants and consequences of causal attributions. *Journal of Personality and social Psychology, 22*. 171-193.

McClellan, B. E. (1992). *Schools and the shaping of character: Moral education in America, 1607-present.* Bloomington: Indiana University: ERIC Clearinghouse.

McClellan, B. E. (1999). *Moral education in America: Schools and the shaping of character from colonial times to the present.* NY: Teachers College Press.

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research 95 (4),* pp.203-215.

Medway, F.J. (1979). Causal attributions for school-related problems: Teacher perceptions and teacher feedback. *Journal of Educational Psychology, 71*, 809-818.

Merriam, S. B. (1988). *The case study research in education.* San Francisco: Jossey-Bass.

Meyer, W. U. (1979). Academic expectations, attributed responsibility, and teachers' reinforcement behavior: A comment on Cooper and Baron, with some additional data. *Journal of Educational Psychology, 71*, 269-273.

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review, 80*, 252-283.

Natriello, G. & Dornbusch, S. M., (1984). *Teacher evaluative standards and student effort.* Research in Teaching Monograph Series. NY: Longman.

Noddings, N. (1984). *Caring: A feminist approach to ethics and moral education.* Berkeley: University of California Press.

Noddings, N. (1992). *The challenge to care in schools: And alternative approach to education.* NY: Teachers College Press.

Noddings, N. (1995). A morally defensible mission for schools of the 21st century. *Phi Delta Kappan, January, 1995:* 365-368.

Noddings, N. (2002). *Educating moral people: A caring alternative to character education.* NY: Teachers College Press.

Odell, C. W. (1925). High school marking systems. *School Review, 33*, 346-354.

Parsons, T. (1959). The school class as a social system: Some of its functions in American society. *Harvard Educational Review, 29,* 297-318.

Peterson, P. L. & Barger, S. A. (1985). Attribution Theory and Teacher Expectancy. In Dusek, J., ed., *Teacher Expectancies*. Hillsdale, NJ: Lawrence Erlbaum, pp. 159-184.

Piaget, J. (1932). *The moral judgment of the child.* NY: Free Press.

Pressey, S. L. (1925). Fundamental misconceptions involved in current marking systems. *School and Society, 21*, 736-738.

Raths, L., Harmin, M., & Simon, S. (1966). *Values and teaching*. Columbus, OH: Merrill.

Rawls, J. (1971). *The theory of justice*. Cambridge: Harvard University Press.

Rejeski, W. J., & McCook, W. (1980). Individual differences in professional teachers' attributions for children's performance outcomes. *Psychological Reports, 46,* 1159-1163.

Rist, R. C. (1970). Student social class and teacher expectations: The self-fulfilling prophecy in ghetto education. *Harvard Educational Review, 40,* 411-451.

Rivers, T. M. (2004). Ten essentials for character education. *The Journal of General Education, 53, (3-4),* 247-260.

Rosenthal, R. & Jacobson, L. (1968). *Pygmalion in the classroom: Self-fulfilling prophecies and teacher expectations.* NY: Holt, Rhinehart, and Winston.

Rothny, W. M. J. (1955). *Evaluating and reporting pupil progress.* Washington, D. C.: National Education Association, p.13.

Rotter, J. B., (1967). Beliefs, social attitudes, and behavior: A social learning analysis. In R. Jesor and S. Feschbach (Eds.), *Cognition, personality, and clinical psychology*. San Francisco: Jossey-Bass.

Ryan, G. W. & Bernard, H. R. (2000) Data management and analysis methods. In N. K. Denzin & Y.S. Lincoln, eds*., Handbook of Qualitative Research, 2ⁿᵈ edition.* Thousand Oaks, CA: Sage, pp. 769-802.

Ryan, K. & Bohlin, K. (1999). *Building character in schools: Practical ways to bring moral instruction to life.* San Francisco: Jossey-Bass.

Sarbin, T. & Allen, V., (1968). Increasing participation in a natural group setting: A preliminary report. *Psychological Record, 18*, 1-7.

Schrag, F. (2001). From here to equality: Grading policies for egalitarians. *Educational Theory, 51*, 63-74.

Schweder, R. (1982). Review of Lawrence Kohlberg's "Essays on moral development. Vol. 1: The philosophy of moral development." *Contemporary Psychology, 27 (2)*, 421-424.

S. G. B. (1840) Weekly reports in Schools. *The Common School Journal, 2, 185-187.* Reprinted in Laska, J. A. & Juarez, T. (1992). *Grading and marking in American schools: Two centuries of debate,* pp. 11-14. In Brookhart, S. M. (2004) *Grading.* Upper Saddle River, NJ: Pearson Education, pp. 18-19.

Sharp, F. C. (1917, 1927). *Education for character: Moral training in the school and home.* Indianapolis, IN: The Bobbs-Merill Company.

Silverstein, M. (1978). An attributional analysis of teachers' evaluative judgments. (Doctoral Dissertation, University of Rhode Island). *Dissertation Abstracts International, 39,* 1055-1056C.

Simon, S., Howe, L., & Kirschenbaum, H. (1972). *Values clarification.* NY: Hart.

Simpson, E. L. (1974). Moral development research: A case study of cultural bias. *Human Development, 17,* 81-106.

Sims, V. M. (1932). Reducing the variability of essay examination marks through the eliminating of variation in standards of grading. *Journal of Educational Research, 26,* 637-647.

Sinclair, R. L. & Ghory, W. I. (1987). Reaching marginal students: A primary concern for school renewal. Berkeley, CA: McCutchan.

Sizer, T. R., & Sizer, N. F. (1999). *The students are watching.* Boston: Beacon Press.

Sockett, H. (1993*).* The moral base for teacher professionalism. New York: Teachers College Press.

Stake, R. E. (1994). *Case studies*. In N. K. Denzin & Y.S. Lincoln, eds*., Handbook of Qualitative Research.* Thousand Oaks, CA: Sage, pp. 236-246.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360-407.

Starch, D. & Elliot, S. C. (1912). Reliability of the grading of high school work in English. *School Review, 20,* 442-457.

Starch, D. & Elliot, S. C. (1913a). Reliability of grading work in mathematics. *School review, 12,* 254-295.

Starch, D. & Elliot, S. C. (1913b). Reliability of grading work in history. *School Review 21,* 676-681.

Stiggins, R. J. (2001). *Student-involved classroom assessment*. Third edition. Upper Saddle River, NJ: Merrill/Prentice Hall.

Stiggins, R. J. & Conklin, N. F. (1992). *In teachers' hands: investigating the practices of classroom assessment.* NY: State University of New York Press.

Strauss, A. & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y.S. Lincoln, eds*., Handbook of Qualitative Research.* Thousand Oaks, CA: Sage, pp. 273-285.

Sullivan, E. V. (1977). A study of Kohlberg's structural theory of moral development: A critique of liberal social science ideology. *Human Development, 20,* 231-242.

Taylor, S. J. & Bogdan, R. C. (1984). *Introduction to Qualitative research and methods: The search for meaning.* New York: Wiley.

Temple University (1968). *Report of the college of education ad hoc committee on grading systems.* pp. 41-48.

Thompson, W. N. (1955). A study of the grading practices of thirty-one instructors in freshman English. *Journal of Educational Measurement, 49*, 65-68.

Thorndike, E. L. (1906). The opportunity of the high schools. *The Bookman*, October, 180-184.

Thorndike, E. L. (1924). Mental discipline in high school studies. *The Journal of Educational Psychology, 15,* 1-22, 83-98.

Thorndike, E. L. & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *The Psychological Review, 18,* 247-261.

Thorndike, R. (1968). Review of Pygmalion in the classroom. *American Educational Research Journal, 5,* 708-711.

Tieg, E. W. (1952). *Educational Diagnosis.* Educational Bulletin #18. Monterey, CA: California Testing Bureau.

Tolman, E. C. (1932). *Purposive behavior in animals and men.* NY: Appleton,-Century-Crofts.

Tom, A. (1984). The moral basis of teaching. In *Teaching as a Moral Craft.* San Francisco: Jossey-Bass.

Tomlinson, A. T. (1994). The easy lie in educating kids. *Education Digest, 60 (2)*: 29-33.

University of California at Berkeley (1965). *Report on methods of evaluating students at the University of California-Berkeley.* October, p.13.

Viadero, D. (2003). Studies illuminate self-defeating behavior by students. *Education Week, 22 (28),* p.8.

Weiner, B. (1980). *Human motivation.* NY: Holt, Rhinehart, & Winston.

Weiner, B., & Kukla, A. (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology, 15*, 1-20.

Weiner, B., Frieze, I. H., Kukla, A, et al., (1971). *Perceiving the causes of success and failure.* Morristown, NJ: General Learning Press.

Weinstein, R. S. (1976). Reading group membership in fifth grade: Teacher behaviors and pupil experience over time. *Journal of Educational Psychology, 68,* 103-116.

West, C.& Anderson, T. (1976). The question of preponderant causation in teacher expectancy research. *Review of Educational Research, 46*, 185-213.

Wiggins, G. (1994). Toward better report cards. *Educational Leadership, Oct.,* pp. 28-37.

Wiggins, G. (1997). Practicing what we preach in designing authentic assessments *Educational Leadership, Dec/Jan.* pp.18-25.

Wood, L. A. (1994). An unintended impact of one grading practice. *Urban Education, 29 (2):* 188-202.

Wynne, E. A. (1997). For-character education. In *The construction of children's character*, A. Molnar, Ed. Chicago: National Society for the Study of Education, pp. 63-76.

Wynne, E. A., & Ryan, K. (1997). *Reclaiming our schools: Teaching character, academics, and discipline,* 2nd ed. Upper Saddle River, NJ: Prentice-Hall.

Yin, R. K. (1989). *Case Study Research: Design and methods,* 2nd edition. Newbury Park. CA: Sage

Zuroff, D. C. & Rotter, J. B. (1985). A History of the Expectancy Construct in Psychology. In *Teacher Expectancies*, J. B. Dusek, ed., Hillsdale, NJ: Lawrence Erlbaum, pp. 9-36.

APPENDIX 1: REPRESENTATIVE INTERVIEW QUESTIONS

*The following is a representative sample of the questions to be asked in the **initial interview** used for collecting data for this study. These questions would serve as springboards for follow-up questions.*

How long have you been an English teacher? Which grade level do you teach? Are you comfortable with this grade level? How long have you taught this grade level?

At the beginning of the school year, do you explain to your classes how you assign grades, and what they mean? Can you tell me what you tell your classes about grades and grading?

Are all your students performing at about grade level? How do you grade those who aren't at the level of the others? Do you grade less able students by a different standard? What about the more able students?

What kinds of assignments will you give this year? How will you use objective tests and more subjective assignments in determining your grades? Does the balance of subjective/objective grading change over the year? If so, how? If not, how do you maintain the balance from one unit to another?

How do you go about grading "creative" assignments? Do you take the individual student's past performance into account in new assignments? What do you do with an assignment that is a "disappointment" to you?

Do your grades serve as a means of communication between you and the students? How? Do you think your students understand what messages you are trying to send through your grades?

Do your grades reflect any kind of judgments about the students as people? That is, do "good kids" get better grades on some assignments than "bad kids?" Do you separate the students' identities from their work, or can you? Should you? How do you see your role as an English teacher in the overall development of your students?

*The following is a representative sample of the questions to be asked in the **mid-quarter marking period interview** used for collecting data for this study. These questions would serve as springboards for follow-up questions.*

You've reached the "five-week" marking period and it's time to send the mid-quarter warnings to students and their parents. What kinds of issues are you weighing now?

You send a "comment sheet" home now. How many of these comments are directly related to grades, and how many address behavior? How do you assess "satisfactory" or "unsatisfactory" behavior? Does the kind of behavior you observe on the part of a student influence your grading in any way? If so, can you describe how? How do you balance behavior into the grading, if you do? Do you worry about fairness?

Have you seen progress in your students so far this year? How will your comments on the five week report communicate your observations? How do you feel about the five week report, especially the first one in the year?

Do parents and students respond to these reports? How? Do you feel the reports fulfill their objective? What value do they have for you?

*The following is a representative sample of the questions to be asked in the **first marking period interview** used for collecting data for this study. These questions would serve as springboards for follow-up questions.*

The first quarter: now you've made some concrete decisions about your students' performance and abilities. How do things stand?

Do you make predictions at this point about how individual students will do this year? What kinds of messages are you trying to send via the grades? What kind of power, if any, do you feel when you submit the grades?

How do you feel about the grades you're giving? Are any of them a little high or a little low? If so, why? Have you given anyone a little push in one direction or the other? Why? Have you done the right thing in each given individual's case? Any reservations? Explain.

Have you faced any dilemmas in assigning grades to anyone? Can you describe them? Any pressures from "outside forces" (administration, coaches, parents, department head, etc.)?

*The following is a representative sample of the questions to be asked in the **midterm interview** used for collecting data for this study. These questions would serve as springboards for follow-up questions.*

What kind of assessments are you making at the half-year mark? Have you made any changes in the way you grade? Are you getting the results you want from your students? Do you feel your grades are sending the messages you want them to?

If you compare grades among students now, do you perceive any that raise moral issues? Is grading a fair process for all involved? Is it fair to you?

Are there some students who are now mathematically doomed to fail? How do you feel about that? Do you have any means of addressing this issue? Do you feel that your power is greater or less than it seemed earlier this year? Explain.

Have any of the issues that arose earlier this year been resolved? Were the resolutions satisfactory to you?

*The following is a representative sample of the questions to be asked in the **final report card grading (for half-year courses) interview** used for collecting data for this study. These questions would serve as springboards for follow-up questions.*

Did you have any failures this year? Any disappointments? Any triumphs? Were there any students who failed that you felt could have passed? Do you feel all your grades were fair?

You have the power to pass or fail a student who is "on the fence." When do you pass such a student, and why? When do you fail such a student, and why? Do you worry about using this power fairly?

How accurate do you feel your grades have been this year? Have the grades sent messages over and above academic performance indicators? Can you give examples where your grades carry a meaning about the student as a person—a lesson you want them to learn about themselves? Do you feel comfortable making decisions about what grades you give? Explain.

Would you rate the course as a success, based on the grades? Do you use grades to judge the success of a course? If so, how?

APPENDIX 2: STUDY INFORMATION SHEET

Study # 01-4525

Indiana University – Bloomington
Study Information Sheet
Moral Dimensions of Grading in High School English

You are invited to participate in a research study. The purpose of this study is to inquire into the dynamics of grading in the high school English curriculum: how teachers perceive grading in terms of purpose, usefulness, and value as a means of communication between teachers and pupils. How you assign grades, weigh their relative values, and explain their meaning to your pupils will be the central focus of this study.

INFORMATION

Collecting the information needed for this study will require between three and five interviews of between 30 and 60 minutes length for each. These will be arranged at your convenience, but should be scheduled at times when grades are a topic of general discussion among English department members: the beginning of the school year, at the mid-way progress report, at the end of the marking period(s), and at the mid-term and final grading period.

This study will involve between ten and twelve teachers. Notes will be taken during the interviews, and a request will be made to tape the interview (to insure accurate and complete recording of the information) before an interview takes place. Taping will only take place with your consent, and you are under no obligation to agree to be taped. Tapes will be destroyed upon completion of the study.

BENEFITS

This study may contribute to an understanding of the complexity of grading in high school English. It may also provide insight into the moral issues teachers weigh out in assigning grades to individual students, and may result in useful reflection upon the meaning and value of grades in English education.

CONFIDENTIALITY

All information will be confidential. Your confidentiality will be insured through the use of pseudonyms and only general descriptors (for example, "a female teacher with five years' experience in teaching tenth grade English") in the report of the study. Neither the school district nor the school will be identified.

CONTACT

If you have questions at any time about the study or the procedures, you may contact the researcher, Laurence Zoeckler, at 1018 Howard Ave., Utica, NY 13501, and (315) 724-7681, or by e-mail at lzoeckle@indiana.edu.

If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have not been honored during the course of this project, you may contact the office for the Human Subjects Committee, Bryan Hall 110, Indiana University, Bloomington, Indiana 47405, by phone at (812)-855-3067, or by e-mail at iub_hsc@indiana.edu.

PARTICIPATION

Your participation in this study is voluntary; you may decline to participate without penalty. If you decide to participate, you may withdraw from the study at any time without penalty and without loss of benefits to which you are otherwise entitled. If you withdraw from the study before data collection is completed your data will be returned to you or destroyed.

APPENDIX 3: MELLMAX HIGH SCHOOL GRADE REPORT FORM

APPENDIX 4.1: SAMPLE COMMENTS FOR THE MELLMAX HIGH SCHOOL GRADE REPORT FORM

11011
6151

```
---------  .....        ------------------------COMMENT------.............................
CODE                                   DESCRIPTION

 028   ATTEMPTS TO ORGANIZE WORK

 029   SELDOM CONTRIBUTES

 030   STUDENT GIVES AVERAGE EFFORT

 031   PUTS FORTH EXTRA EFFORT

 032   WORKING VERY HARD

 033   HAS SHOWN IMPROVEMENT

 034   HAS SHOWN SIGNIFICANT IMPROVEMENT

 035   WORKING TO HIS/HER POTENTIAL

 036   CONSIDERATE OF OTHERS

 037   CONSCIENTIOUS AND HARDWORKING

 038

 039   PARTICIPATES IN CLASS

 040   SHOWS GOOD WORK HABITS

 041   NEEDS TO BUILD CONFIDENCE

 042   ATTENTIVE IN CLASS

 043   PARENT CONFERENCE NEEDED, PLEASE CONTACT ME

 044   PLEASE UTILIZE THE HOMEWORK HOTLINE - 894-2657

 045   STRUGGLES WITH MATERIAL, BUT TRIES HARD

 046   COMES FOR EXTRA HELP

 047   PLEASANT

 048   WELL-BEHAVED

 049   NOT WORKING TO POTENTIAL

 050   FREQUENT ABSENCES IMPAIR UNDERSTANDING OF MATERIAL

 051   WORKS WELL IN A GROUP

 052   COMES TO CLASS WITHOUT MATERIALS OR ASSIGNMENTS

 053   DOES NOT KEEP TRACK OF ASSIGNMENTS

 054   DOES NOT COMPLETE REQUIRED WORK
```

254

APPENDIX 4.2: SAMPLE COMMENTS FOR THE MELLMAX HIGH SCHOOL GRADE REPORT FORM

```
-----------------------------------------COMMENT----------------- -------------------------
CODE                                    DESCRIPTION

303  STUDENT WORKS WELL IN A GROUP
304  STUDENT DOES NOT TAKE THE OPPORTUNITY TO IMPROVE GRADE(S).

305  QUIZ AND TEST GRADES ARE POOR

306  QUALITY OF STUDENT'S WORK NEEDS IMPROVEMENT

307  FREQUENT ABSENCES IMPAIR UNDERSTANDING OF SUBJECT

308  STUDENT NEEDS TO STUDY MORE

309  STUDENT FAILS TO COMPLETE HOMEWORK ASSIGNMENTS

310  STUDENT'S HUMOR IS APPRECIATED

311  STUDENT FAILED TO TURN IN FINAL EXAM/PROJECT ON TIME

312  STUDENT FAILED TO TURN IN A MAJOR PROJECT ON TIME

313  STUDENT NEEDS TO SEE ME DURING ADVISEMENT OR AFTER SCHOOL

314  HOPE TO SEE STUDENT IN ADVANCE HEALTH CLASS

315  I WILL MISS STUDENT NEXT SEMESTER

320  CONTRIBUTES TOWARD A POSITIVE CLASSROOM ATMOSPHERE

321  NEEDS TO DEVELOP AND MAINTAIN A MORE POSITIVE ATTITUDE

322  TAKES ADVANTAGE OF EXTRA CREDIT ACTIVITIES

323  SUCCESSFUL LAB WORK HAS RAISED OVERALL GRADE

324  A BETTER UNDERSTANDING OF LAB WORK WILL IMPROVE OVERALL GRADE

325  SHOULD SEEK MORE TEACHER HELP OR PEER HELP IN UNDERSTANDING AREAS BEING TESTED

326  STUDENT HAS NOT CONSISTENTLY TURNED IN ESSAY(S) WEAK MISTAKES NEED CORRECTION.

327  LOW GRADES ARE THE RESULT OF NOT RETAKING QUIZ/TEST WHEN THE OPPORTUNITY OFFERED

328  STUDENT NEEDS MORE HELP WITH WRITING SKILLS

329  STUDENT COMPLETED AN EXCELLENT PROJECT THIS QUARTER

330  NEEDS TO IMPROVE LAB WORK AND MEET DEADLINES

331  SHOWS CONCERN FOR CLASSMATES

332  COMPLETES LABS ON TIME

333  WILLING TO HELP OTHERS
```

APPENDIX 4.3: SAMPLE COMMENTS FOR THE MELLMAX HIGH SCHOOL GRADE REPORT FORM

MARK REPORTING MASTER LISTING

TEACHER COMMENT FILE

```
-------------------------------------COMMENT-------------------------------------
CODE                              DESCRIPTION

334  SELF-ESTEEM PROJECT WAS TERRIFIC

335  FITNESS PROJECT WAS OUTSTANDING

336  SUBSTANCE PROJECT WAS VERY CREATIVE

337  NEEDS TO COMPLETE WORK WEEKLY TO BETTER UNDERSTAND TOPIC

338  WILLING TO GET HELP FROM PEERS

339  STUDENT'S DUTY SCHEDULE MAKES IT DIFFICULT TO OBTAIN EXTRA HELP

340  STUDENT IS DOING SATISFACTORY WORK

343  STUDENT'S RECENT WORK HAS BEEN DECLINING

345  POSITIVE PARTICIPATION IN CLASS

346  GENUINELY ATTENTIVE IN CLASS

353  NEEDS TO BE MORE ORGANIZED

354  NEEDS WORK ON TAKING NOTES

355  NEEDS TO EXPAND ATTENTION SPAN

358  WILL IMPROVE BY COMPLETING ASSIGNMENTS

360  STUDENT'S RECENT WORK HAS BEEN IMPROVING

361  STRUGGLES WITH MATERIAL, BUT TRIES HARD

362  COMES FOR EXTRA HELP

363  ALTHOUGH DOING SATISFACTORY WORK, IS CAPABLE OF DOING BETTER

365  STUDENT'S WRITING SKILLS NEED IMPROVEMENT

366  ENGLISH COMPREHENSION IS POOR

368  STUDENT NEEDS TO PARTICIPATE MORE IN CLASS

370  CLASSROOM BEHAVIOR NEEDS TO BE MORE POSITIVE

371  DOES NOT EXHIBIT PROPER RESPECT FOR AUTHORITY

372  DISRUPTS CLASS

373  OCCASIONAL NEEDS TO LISTEN AND PAY MORE ATTENTION IN CLASS

380  STUDENT NEEDS TO SEE ME FOR EXTRA HELP
```

APPENDIX 4.4: END-OF-YEAR DIRECTIONS TO TEACHERS FOR THE MELLMAX HIGH SCHOOL GRADE REPORT FORM

MEMORANDUM

TO: Classroom Teachers
FROM: Director of Guidance
RE: MARK REPORTING FORMS – Report Cards
DATE:

Attached please find the END OF THE YEAR Mark Reporting forms for your classes. When completing these forms, please follow these procedures:

1.) Use No. 2 pencils only
2.) Make heavy black marks that fill the circle completely.
3.) Erase gently and completely any marks you wish to change.
4.) Make no stray marks.
5.) Complete the Mark Reporting Form for each student.

    A.) In the **PREVIOUS MARK** area, write the student's grade and absences.
    B.) In the **CURRENT PERIOD MARK** area, fill the appropriate circles.
    C.) Do Not Use the **CURRENT PERIOD EXAM** area.
    D.) In the **COMMENT 1** area, fill in circles 1, 2, 3, or 4 for student attitude.

        1 = SUPERIOR EFFORT SHOWN IN CLASS
        2 = EXCELLENT EFFORT SHOWN IN CLASS
        3 = SATISFACTORY EFFORT SHOWN IN CLASS
        4 = IMPROVEMENT NEEDED IN CLASSROOM EFFORT

    E.) In the **COMMENT 2** area, fill in circles 5, 6, 7, or 8 for student behavior.
        5 = SUPERIOR BEHAVIOR DISPLAYED IN CLASS
        6 = EXCELLENT BEHAVIOR DISPLAYED IN CLASS
        7 = SATISFACTORY BEHAVIOR DISPLAYED IN CLASS
        8 = IMPROVEMENT NEEDED IN CLASSROOM BEHAVIOR
    F.) The **COMMENT 3** area may be used, if you wish.
    G.) In the **CLASS ABSENCES** area, fill in the appropriate circle. {REQUIRED}
    H.) The **FINAL EXAM** area should contain class Final.
    I.) The **FINAL AVERAGE** area should contain the Final School Grade.

6.) **DEADLINE** Please submit all completed Class forms to your Department Chairperson by 12:00 P.M. on Monday, June 20, OR BEFORE. Department Chairs will check them for accuracy and submit them to   by 12:00 p.m. Tuesday, June 21. Regents Exam Report Forms must be to   by 3:00 p.m., June 24.

    A.) Double-check your forms for accuracy before submitting them.
    B.) Be certain to return all pages in page number order.
    C.) Please use Post-It Notes to indicate any changes on the page, be sure they stick out beyond the edge of the page.
    D.) If a student is NOT on your forms, please check with one of the counselors in order to verify enrollment. DO NOT "Red" line any students.

If you have any questions, please see your Department Chairperson.

APPENDIX 5.1: SAMPLE GRADING RUBRIC FROM MELLMAX HIGH SCHOOL

Grading Criteria:

1 poor   2-weak   3-fair   3.5-average   4-above average   4.5 good   5-excellent

Quality and development of ideas	1 2 3 3.5 4 4.5 5	
Organization, relevance, movement	1 2 3 3.5 4 4.5 5	
Subtotal		X 6 = _____

Style, flavor, individuality	1 2 3 3.5 4 4.5 5	
Wording, phrasing	1 2 3 3.5 4 4.5 5	
Subtotal		X 1 = _____

Grammar, sentence structure	1 2 3 3.5 4 4.5 5	
Punctuation/capitalization	1 2 3 3.5 4 4.5 5	
Spelling	1 2 3 3.5 4 4.5 5	
Manuscript form, legibility	1 2 3 3.5 4 4.5 5	
Subtotal		X 1.5 = _____

Total Score _____

APPENDIX 5.2: SAMPLE GRADING RUBRIC FROM MELLMAX HIGH SCHOOL

Name_____   Date_____
English 11 _____   RESEARCH PAPER

RESEARCH PAPER RUBRIC

TOPIC:_____ ...   _____

Points earned   Category

Organization and development (30 points)

Style/Wording (5 points)

Mechanics (10 points)

Citation (5 points)

Total points earned: /50

APPENDIX 5.3: SAMPLE GRADING RUBRIC FROM MELLMAX HIGH SCHOOL

Name.
English 10 _____

10th grade research paper

10th GRADE RESEARCH PAPER: ROUGH DRAFT-SCORING RUBRIC

QUALITY	Responses at this level: 5	Responses at this level: 4	Responses at this level: 3	Responses at this level: 2	Responses at this level: 1	Responses at this level: 0
Meaning: the extent to which the response exhibits sound understanding, interpretations, and analysis of the task and text(s)	reveal an in-depth analysis of the documents / makes insightful connect-ions between in-formation and ideas in the documents and the assigned task	convey a thorough understanding of the documents / makes clear and explicit connections between information and ideas in the documents and the assigned task	convey a basic understanding of the documents / make implicit connections between information and ideas in the documents and the assigned task	convey a basic understanding of the documents / makes few or superficial connections between information and ideas in the documents and the assigned task	convey a confused or inaccurate under-standing of the documents / allude to the documents but make unclear or unwarranted connections to the assigned task	provide minimal or no evidence of understanding / make no connections between information in the document and the assigned task
Development: the extent to which ideas are elaborated using specific and relevant evidence from the document(s)	develop ideas clearly and fully, making effective use of a wide range of relevant and specific details from the documents / meets the required number of pages and sources	develop ideas clearly and consistently, using relevant and specific details from the documents / missing one page or one source	develop some ideas more fully than others, using specific and relevant details from the documents / missing 2 pages or 2 sources	develop some ideas briefly, using some details from the documents / missing 3 pages or 3 sources	are incomplete or largely undeveloped, hinting at ideas, but references to the documents are vague, irrelevant, repetitive, or unjustified / missing 4 pages or 4 sources	are minimal, with no evidence of development / missing 5 or more pages or sources
Organization: the extent to which the response exhibits direction, shape, and coherence	maintain a clear and appropriate focus / exhibit a logical and coherent structure through skillful use of appropriate devices and transitions / support a clear and appropriate thesis	maintain a clear and appropriate focus / exhibit a logical sequence of ideas through the use of appropriate devices and transitions / a clear thesis is fairly well supported	maintain a clear and appropriate focus / exhibit a logical sequence of ideas but may lack internal consistency / thesis is somewhat supported	Establish, but fail to maintain, an appropriate focus / exhibit a rudimentary structure but may include some inconsistencies or irrelevancies / attempt to support thesis but lose focus	lack an appropriate focus but suggest some organization, or suggest a focus but lack organization / vague thesis or virtually no research based support	Show no focus or organization / no support for thesis or thesis is not evident

10th GRADE RESEARCH PAPER: ROUGH DRAFT-SCORING RUBRIC

QUALITY	Responses at this level: 5	Responses at this level: 4	Responses at this level: 3	Responses at this level: 2	Responses at this level: 1	Responses at this level: 0
Language Use / Conventions: the extent to which the response reveals an awareness of audience and purpose through effective use of words, sentence structure, and sentence variety / mechanics (spelling, punctuation, paragraphing, capitalization, and grammar)	are stylistically sophisticated, using language that is precise and engaging, with a noticeable sense of voice and awareness of audience and purpose / vary sentence structure and length of sentences to enhance meaning / demonstrate control of conventions with essentially no errors	use language that is fluent and original, with evident awareness of audience and purpose / vary structure and length of sentences to control rhythm and pacing / demonstrate control of conventions, exhibiting essentially occasional errors	use appropriate language, with some awareness of audience and purpose / occasionally make effective use of sentence structure or length / demonstrate partial control, exhibiting occasional errors that do not hinder comprehension	rely on basic vocabulary, with little awareness of audience or purpose / exhibit some attempt to vary structure or length for effect, but with uneven success / demonstrate emerging control, exhibiting occasional errors that hinder comprehension	use language that is imprecise or unsuitable for the audience or purpose / reveal little awareness of how to use sentences to achieve an effect / demonstrate a lack of control, exhibiting frequent errors that make comprehension difficult	are minimal / use language that is predominantly incoherent, inappropriate, or copied directly from the text / making assessment of conventions unreliable, may be illegible
Citation: the extent to which the response is cited in accordance with MLA standards	in-text citation and works cited page follow MLA format completely / in-text citations accurately match works cited page	in-text citation and works cited page exhibit few or minor errors / very few errors in matching citations to works cited page	in-text citation and works cited page exhibit occasional errors but knowledge of format is evident / some errors in matching citations to works cited page	in-text citation and works cited page exhibit frequent errors but some knowledge of format is evident / frequent errors in matching citations to works cited page	in-text citation and works cited page exhibit excessive errors, minimal knowledge of format / excessive matching problems but some attempt was made	no works cited page or no knowledge of format evident / works cited page does not match any in-text citation / in-text citation is not present

Vita
Laurence G. Zoeckler

EDUCATION AND PROFESSIONAL TRAINING

Doctor of Philosophy, Curriculum and Instruction, **2005**
School of Education, Indiana University-Bloomington, IN
Master of Arts in English Literature, **1987**
Syracuse University, Syracuse, NY
Master of Arts in Religious Studies, **1986**
St. Joseph's Seminary, Yonkers, NY
Bachelor of Arts, magna cum laude, **1980**
Hamilton College, Clinton, NY

INSTRUCTIONAL EXPERIENCE

Assistant Professor of Education, Utica College of Syracuse University, 2001-present.
- *Courses include Secondary Methods for English and Social Studies, Foundations of American Schools, Graduate Research Methods for Education, Student teaching seminar, Apprentice teacher mentoring*
- *Interim Coordinator of Student Teaching, spring 2002 & 2005*
- *Supervise Apprentice Teachers*

English teacher, Ilion Junior-Senior High School, 1988-2001, English Dept., Ilion High School, Ilion, NY
- *English courses in all grades from 8-12*

Adjunct Instructor in Education, **1988-1996** &**1999-2001**
Education Department, Utica College, Utica, NY
- *Courses have included Methods and Curriculum in Teaching, Reading and Study in the Secondary School, and Practical Teaching Methodologies*
- *Promoted to Adjunct Professor of Education in 1993*

Associate Instructor, Indiana University, **1996-1998**
Office of Student Teaching, School of Education, Bloomington, IN
- *Supervised student teachers at all levels, K-12, in Indiana public schools; Conducted seminars for student teachers*

Teaching Assistant, Syracuse University, **1986-1987**
English Department, Syracuse University, Syracuse, NY
- *Rhetoric and Composition*

ACADEMIC AWARDS AND HONORS

University Fellowship, Indiana University-Bloomington
renewed each semester, **Fall 1996-Spring 1998**
University Scholarship, Indiana University, **1996-1998**
University Assistantship, Indiana University, **1996-1998**
University Assistantship, Syracuse University**, 1986-1987**
Duell German Prize, Hamilton College, **1980**
Departmental Graduation Honors, German, Hamilton College**, 1980**
Inducted into Phi Sigma Iota, National Foreign Language Honor, **1979**