

MORAL ASPECTS OF GRADING: A STUDY OF HIGH SCHOOL ENGLISH TEACHERS' PERCEPTIONS

AUTHOR

LAURENCE G. ZOECKLER is Assistant Professor of Education at Utica College in Utica, NY.

ABSTRACT

Moral issues embedded in grading practices of high school English teachers are central to this study. Arriving at a fair grade while weighing both achievement and non-achievement factors and the role of teachers' expectations are examined using a theoretical framework which considers grading processes in terms of truth, worthwhileness, trust, and intellectual and moral attentiveness. Grading is influenced by grading systems, perceptions of effort, and concern for moral development. Results indicate that English teachers struggle with issues of fairness, but are confident that their grades communicate the messages they hope to send.

INTRODUCTION

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 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 English classes, 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 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 response. A creative writing assignment may receive an "A" from a teacher in one class, but a "C" from a teacher in another class. Thus, English classes pose a special problem when converting student performance to grades. 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. The moral issues surrounding grading are perhaps nowhere as evident as they are in English courses.

CONTEXT AND LITERATURE

Since the turn of the century, grading and giving marks in school have been a source of controversy (Cureton, 1971; Ebel & Frisbie, 1986; Hopkins, Hopkins, & Stanley, 1990). Ebel and Frisbie (1986) 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.

Feldman, Alibrandi, and Krkopf (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". They asked "How do teachers decide what grade to put on students' report cards?"(p.141) and produced a laundry list of devices teachers use to amass points and calculate averages. However, they did not explicitly consider the moral issues involved in grading.

Hendrickson and Gable (1997) lamented 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; clearly, grading is heavily laden with moral considerations that go well beyond the requirements of "achievement testing."

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 have argued that the current culture is one which enforces a “values-neutral” stance upon the schools and results in a system where teachers no longer know what their role in moral education is (Delattre & Russell, 1993; Gecan & Mulholland-Glaze, 1993), many others have argued that schooling is by its very nature a moral undertaking (Jackson, Boostrom, & Hansen, 1993; Sockett, 1993; Goodlad, Sirotnik, & Soder, 1990). 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. Dewey (1909) argued that schoolteachers taught morals “every moment of the day, five days a week” (p.3).

This indirect approach to moral education nevertheless 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) explained 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). Rawls (1971) asserted that “A well-ordered society both advances the good of its members and is regulated by a public conception of justice” (p. 23). According to Kerr (1987) “Education 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.

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” that tells nothing about their actual performance (Tomlinson, 1994). Grade depression (Wood, 1994) or grade inflation (Bracey, 1994) may creep into school practices as parents—and state education departments—demand improved performance and grades seem the only way to measure it. Dockery (1995) pointed out that grading scales are often arbitrary and vary from teacher to teacher, that grades may be used to influence behavior and thus not accurately depict academic performance, and “zeros are motivation killers” (p. 34).

Teacher expectations and perceptions of student attitudes also affect grades. Rosenthal and Jacobson's (1968) landmark study of teacher expectations supported the notion of the self-fulfilling prophecy by suggesting that teachers often judged their students based more on expected performance than on actual performance. Kirschenbaum, Napier, and Simon (1971) worried that teacher expectations of student performance could be based on little more than word of mouth. Closer examination of the Rosenthal and Jacobson experiment suggested weaknesses in the research design (Barber, 1969; Elashoff & Snow, 1971; Jensen, 1969; Thorndike, 1968). Teacher expectations and their relationship to grading have been a topic of study ever since.

Student effort and attitude are tied to teacher expectations, and all of these are in turn tied to student motivation. Grades are often used as a device for motivating students (Brookhart, 2004; Glasser, 1971; Hargis, 1990; Karmel, 1970). Viadero (2003) researched self-handicapping behaviors and avoidance strategies in high school students, and concluded that teacher expectations based on success rather than attempts at success discouraged effort and encouraged self-defeating behaviors.

RESEARCH QUESTIONS

Brookhart's (1991) critique of prevalent grading practices described 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 whole. In an attempt to explore the "hodgepodge grading," that Brookhart (1994) decried, Cross and Frary (1999) suggested 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 recommended 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). Studies regularly demonstrate that classroom teachers continue to consider multiple factors as well as achievement. It is clear that teachers' understanding of what grades mean and how they are meant to be understood is an area in need of exploration.

This study explored grading practices and issues of judgment, communication, and character development in grading through a frame-

work which exposes the underlying moral issues in grading. The study began with these central questions:

1. *"How do English teachers decide what is right in giving grades to individual students?"*
2. *"How and what do English teachers intend to communicate to students through grades?"*
3. *"Is character development a part of the role of an English teacher, and if so, does this influence grading, and how?"*

FRAMEWORK FOR THE STUDY

A framework for considering the moral issues involved in grading was developed in order to guide the data collection and facilitate the analysis of the data gathered for this study. Drawn chiefly from Jackson et al. (1993), the key elements observed to be at work in the "Moral Life of Schools" study were transferred to the issues of grading and augmented in order to provide a means of "weighing out" the moral implications of grading practices among the participant teachers in this study.

Jackson et al. (1993) determined that the notions of right and wrong, the common good, and desirable ends were implicitly accepted in classrooms and schools. In addition, they found that three other ideas were implicit in the day-to-day operations of schools. These are the notions of "truthfulness," "worthwhileness," and "trust." These were referred to as "substructural" elements of schooling rather than "structural" ones to emphasize their fundamental embeddedness.

"Truthfulness" is taken for granted in what the teacher transmits in class, in what the students hand in as assignments, in grading policies, in communication between students and teachers, and with parents and the community. Without this basic moral feature, schooling would be impossible. The extent to which grades and grading actually account for and reflect the truth about students in a classroom is an important measure of the moral value of both the process and the product of grading.

"Worthwhileness" answers the question, "What are we doing this for?" and the assumption of worthwhileness "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" (Jackson et al., 1993, p.25).

This assumption is “a tacit acknowledgment of the moral character of the institution” (Jackson et al., 1993, p.26). Like truth, worthwhileness in school is also taken for granted, yet every student and most teachers, parents, and administrators sometimes find themselves questioning the worthwhileness of some elements of schooling. Where this affects grades and grading, the moral issue arises: “Is this truly worthwhile, and if not, why are we grading it?”

“Trust” becomes the third substructural element, in part because it 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.

The issue of trust is carried still further by Baier (1995), who explained 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 explores the range of discretion that the subject teachers believe is appropriate and acceptable to their own consciences in their grading practices. It also investigates the extent to which Jackson’s et al (1993) substructural constructs inform the grading process and its outcomes in the work of the teacher participants.

The three elements embedded in the philosophical roots of schooling suffice to give an indication of the moral dimensions of decisions guiding choices of grading and content. Hansen (1997) helped to shed light upon the moral dimensions of both grading and instruction by incorporating two additional concepts, “moral attentiveness” and “intellectual attentiveness.”

Put briefly, 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.

Hansen’s (1997) notion of moral attentiveness, on the other hand, has two features: “alertness to the development of students’ character, and awareness of one’s regard and treatment of students” (p.8). 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.

The framework, then, consists of five elements: truthfulness, worthwhileness, trust, intellectual attentiveness, and moral attentiveness. The moral issues in grading can be analyzed by the extent to which actual practices in grading reflect these five elements.

SAMPLE

The study reported here was a case study involving the English department at Mellmax High School (a pseudonym), a rural high school in upstate New York. 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 constituted a purposive sample that included the entire department of ten teachers, one of whom was male. In addition, two recently retired English teachers were interviewed for purposes of comparison. A thirteenth teacher (a male) who had recently left the district and taken a similar job in another county also participated, providing evidence that grading practices among the participant teachers was not atypical. All teachers were certified to teach English, in accordance with state regulations.

DATA COLLECTION

Interviews with each of the teachers were conducted over the course of an entire school year. The interviews were semi-structured 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 terms of fairness and perceptions of fairness. Initially, moral issues involved in grading were examined through indirect questions in order to keep the participants at ease and elicit frank responses. Several 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 adjustments were made to grading policies and how such adjustments were justified in the minds of the teachers.

Interviews were scheduled during the school day and required approximately an hour each; the interview schedule was adjusted as required; teachers were encouraged to clarify and expand upon answers to questions, to give examples of typical and unusual grading decisions,

and to explain their thinking in making decisions about assigning grades. Each teacher was interviewed at least three times—at the first marking period, the midterm, and the final marking period; several of the teachers were interviewed (dependent upon their availability) at intervals between these periods. Interviews were taped and transcribed for analysis.

While interviews comprised the bulk of the data, documents such as the teachers' handbook, school board policy statements on grading, and grade-books kept by the teachers were examined when they were useful to provide additional information. The high school principal was also interviewed in order to obtain information about grading practices at the high school and within the English department. Field notes were kept along with the interviews, and these notes were expanded as interviews were compared with one another. The notes were taken in order to add information which could not be captured in the recording, e.g., facial expressions which supported or contradicted spoken remarks, body language, and notes to explain references to grade books, bulletin boards, or assignments, and so on. The notes also provided details about the period of the day, whether the teachers appeared tired, anxious, at ease, etc. Some of the notes were made beforehand (e.g., period of the day, time of the school year, appearance of the classroom). Most of the notes were taken during the interview or immediately thereafter. At times, field notes also contained the interviewer's reaction or a commentary on a teacher's remark, either because the remark seemed unusual or because it suggested a new line of questioning worth following up with the other teachers.

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. 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 shared with the participants for validation. 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. Data collected in "feedback interviews" were then used to further delineate the issues involved.

After the data were categorized and verified, a process of inductive analysis was undertaken in order to draw conclusions about the information obtained. Support for conclusions was drawn from quotations taken

directly from the data. Negative case evidence that seemed to contradict the main body of evidence was considered carefully and presented along with the main analysis.

As issues arose, the pertinent literature surrounding the issues was examined and summarized. Current theories that deal with grading practices and their consequences were examined and the practices of the participant teachers were considered in light of these theories. Areas of congruence with, and divergence from, theory were considered in order to understand the grading practices of teachers in the field more fully. The initial analysis yielded six categories or themes: "grading systems," "subjective judgment," "rubrics," "dealing with effort," "forming character," and "fitting to the system." Some themes were folded into others in subsequent analysis, resulting in three major topics:

GRADING SYSTEMS

The ways in which teachers structured their grading, assigned values to different assignments, and weighted categories of similar assignments like homework, tests and quizzes, essays, projects, and so on. This section also examined the wider school-district system of report cards and its effects on the ways teachers synthesized their recorded grades into the single grade required for the report card.

EXPECTATIONS, EFFORT AND ATTITUDE

How teachers formed initial expectations of their students' work and conveyed these expectations in their grading, and how those expectations changed as the school year progressed, causing changes in assignments and in grades. This section also included how such expectations colored teachers' interpretations of student effort and attitude. Finally, this section considered how teacher expectations influenced student expectations of what was required, how it would be graded, and what assignments were worthwhile.

CHARACTER DEVELOPMENT

The extent to which issues of character development influenced grading practices and whether the teachers used grades to influence the development of their students' character.

FINDINGS

GRADING SYSTEMS

All of the teachers used grading systems that made use of points and weights. Each assignment that was graded received some number of

points depending upon its assigned value. Some teachers simply totaled the points and divided by the number of points that were possible; others categorized assignments (e.g., homework, tests & quizzes, oral reports, term papers, etc.) and a certain percentage of the grade was assigned to each category. All of the teachers indicated that the school's report system limited the teachers' means of expressing their evaluations of their students because it allowed for only a single numerical grade and up to three "comments" such as "fails to do homework," "does poorly on tests," or "a pleasure to have in class," from a standard list. Comments were selected by number and "bubbled in" with pencil on a Scantron grading sheet. Several findings emerged when grading systems were considered in terms of the theoretical framework established for this study:

- Using a point system to arrive at numerical averages freezes grades at a given moment. This practice requires that a work in progress be treated as a final product, so that the grade given at any moment using a point system may not reveal the actual truth about a student;
- The demand for grades at fixed time intervals (i.e., marking periods) means teachers may pull together a few assignments, grade them and average their values, present a grade-for-the-time-being that may change with the very next assignment, and be a false one by the time the student actually receives it. **All the participant teachers felt this was a fault in the grading process, especially in the case of the first five-week grade** which came too soon for the teachers to feel that they had any firm sense of their students' performance so early in the year;
- Because the grading system demands fixed-moment grades, it freezes the grades whether or not they give true representations of the students' work, and then averages them with other fixed-moment grades. **Thus, an inaccurate reflection of student work, once "frozen," becomes part of the official final grade – which can be no truer than the grades upon which it was based in the first place.** Poor performance early in the year results in poor grades that must be averaged against subsequent grades – whether or not the student makes phenomenal progress in the interim.
- Weighting of assignments can be manipulated in ways that threaten the truthfulness of grades. Difficult assignments can be discounted and easier ones inflated to offset poor class averages. **The younger teachers, especially the untenured ones, often "fixed" their weighting in order to keep failure rates within a range they believed the admin-**

- stration would find acceptable;
- One teacher's search for "patterns of performance" led her to throw out grades that did not fit a perceived pattern. This meant that her grades may not have told the truth about a student whose performance pattern was beginning to change – whether for better or worse;
 - Heterogeneous grouping means that some assignments are too easy for some students or too hard for others – and thus not equally worthwhile for all students;
 - Unfinished lessons – or lessons not yet mastered – are often left behind when a unit ends, sending the message that what a student did not learn during the unit's fixed time frame can be left behind, and therefore suggesting that the remainder is not as worthwhile as it was when first presented. This also casts doubt on the entire lesson's (or unit's) worthwhileness, since different students have left different parts unlearned;
 - The teachers recognized this fault in moving forward before students had done all they could, and most allowed students to re-do some assignments – often, ironically, permitting that only the worthwhile ones could be re-done;
 - The concept of "worthwhileness" may be partly grounded in the notion of "usefulness" as opposed to "inherent value," making grading based on a system worthwhile if only because it furthers a student's progress through the school system;
 - Trust is essential to building credibility in a teacher's grading system—students must trust their teachers to devise a fair system and calculate grades accordingly. At the same time, teachers must trust their students to make their best efforts and give honest performance on their assignments.

EFFORT, EXPECTATIONS, AND ATTITUDE

At the beginning of the year, all the teachers were asked, "When you grade subjective assignments like essays, do you look at the names of the students?" All but one (the longest serving veteran) said that they did not because they did not want the name to influence their grading. However, when they were asked the same question at the midterm, all said they did look at the names. One realized the contradiction: "You asked me that question before, didn't you? And I think I gave you a different answer." Asked for an explanation, she remarked, "Well, at the beginning of the year, I didn't know my students yet. Now I do, so I know what to expect and what to compare this essay with [i.e., earlier work]." All the other

teachers responded in a similar way. This was in keeping with the findings for this area of the study:

- The participating teachers' remarks indicate that they use grades both to form expectations ("truth finding") and to reveal expectations ("truth-telling"). These dual roles of grading are rooted in truth;
- At the beginning of the year, teacher expectations are based on previous experience. This is tempered as assignments are given and graded and teachers begin to know their students.
- If student performance is false, the expectations drawn from that performance are false; when expectations are false, grades do not tell the truth about a student's abilities;
- Performance different from expectations can undermine a teacher's trust in a student, since it may be the result of cheating (higher performance than expected) or lack of effort (lower performance than expected). Grading on an assumption of either of these, however, can undermine student trust in the teacher, since expectations must take into account actual performance, which can sometimes be better or worse than expected;
- Students' expectations of the course, the teacher, and the grades in the course are founded on an assumption of worthwhileness. When this fails, attitude affects performance, which no longer matches teacher expectations, and grades become an inaccurate reflection of student ability and/or potential;
- The participating teachers indicated that their assumptions regarding the worthwhileness of their assignments influence their expectations, and poor performance is often seen as a reflection of "bad attitude" rather than a reflection of negative student judgment about the worthwhileness of an assignment.
- Both moral and intellectual attentiveness are required for forming and maintaining accurate teacher expectations for student performance and interpretation of attitudes.

CHARACTER DEVELOPMENT AND GRADING

Three different models for character development are common in public schools: "character education," which transmits a list of virtues to be learned; "moral education," which explores and discusses moral issues in order to develop a means of understanding moral matters; and "caring," which focuses on the value of the individual as a person, i.e., in terms of inherent worth. In most schools, all three are used in different places and

under differing circumstances. The Mellmax High teachers used an eclectic approach to moral training, combining features of all three models. All the participating teachers' comments indicated that character development was very important and part of their role as teachers.

Literature was universally perceived as a vehicle for discussion of moral issues and questions about character. But the findings regarding character development and grading in terms of the framework suggest that the relationship is very subtle, and almost unperceived by most of the teachers themselves.

- None of the teachers asserted that character could be graded directly, and all said they avoided incorporating perceptions of student character into their grades. Nevertheless, all admitted that concrete expressions of character (diligence, determination, carelessness, laziness, willingness to help others, etc.) that affected performance did affect grades;
- All the teachers at Mellmax High stated that being a role model for good character was part of their duty, and attempted to model the kinds of behavior they hoped their students would emulate. All reported sometimes consciously drawing their students' attention to the moral aspects of their own (teachers' or students') actions;
- Trust is an important part of any model for character development; however, the level of trust required for character education (the list of virtues approach) is lowest, requiring only that the virtues being taught are appropriate. The need for trust increases with moral education, where the teacher's way of looking at moral matters requires trust in order to be accepted. In the caring approach, the level of trust is deepest because it requires a personal relationship between the one who cares and the one who is cared about;
- Worthwhileness is assumed in character development; this facet of schooling was widely stated to be of primary importance by the English teachers at Mellmax High, but even in this age of measurement madness, this most important piece of learning was seen as ungradeable due to the wide variety of understandings about the moral in our pluralistic society and the lack of agreement about ways to measure it.
- Moral and intellectual attentiveness are essential to the development of students' character; all the English teachers at Mellmax High insisted they were acutely aware of the need to be mindful of their students' development in this area.

- The participating teachers' remarks indicated that their grading did not reflect judgments about their students' character, **except in indirect ways such as passing a borderline student who was judged to be of good character or failing a borderline student who was judged to be of bad character.** However, if this rule guided the grading of many smaller assignments as well as the calculation of marking period averages, the apparently small influence of character judgments made by teachers about their students over the span of a full marking period may have a greater cumulative effect on students' grades than may at first appear.

RESEARCH QUESTIONS REVISITED

Returning to the original research questions, the answers that follow capture the essence of the findings of this study. They should not be accepted as definitive or absolutely comprehensive answers, however. Capturing the entire complexity of moral issues in grading is beyond the grasp of such limited research; a single case study like this one serves only to reveal some of the depth of the moral issues embedded in grading.

HOW DO ENGLISH TEACHERS DECIDE WHAT IS RIGHT IN GIVING GRADES TO INDIVIDUAL STUDENTS?

Teachers arrive at grading schema through structure, purpose, and understanding of the constraints that surround them. The system for grading itself can be a major factor in the way grades are given. **For example, grades are influenced at least in some part by the teacher's understanding of the "acceptable" number of failures in a class, and teachers (especially untenured ones) struggle with this in order to give a fair grade to each student.**

The "right grade" reflects personal philosophy, understanding of the student, and reflection. Some lessons are clearly more important than others, even if the grading scheme cannot make this so, and teachers attempt to impress this on their students when explaining how they arrive at grades. Most teachers struggle to communicate the value of their grades to their students. Both moral attentiveness and intellectual attentiveness demand that teachers tailor grades in very individual ways, despite the fact that their school grading system imposes at least an appearance of uniformity among grades.

HOW AND WHAT DO ENGLISH TEACHERS INTEND TO COMMUNICATE TO STUDENTS THROUGH GRADES?

Teachers attempt to communicate many different things through their grades, all tailored to the individual student. Some of these include: level of expectation, level of academic achievement, encouragement, and disappointment. Often narrative grading is used in addition to the assigning of number or letter grades (which are converted to numbers for averaging for report card purposes). This narrative grading is used to add to the meaning of a number or letter grade, or to encourage the student, especially if the grade is out of keeping with either the teacher's or the student's expectations.

Grading systems and practices are unique to each teacher and shared understanding between students and their teachers about the meaning of grades is constantly sought by the teachers, who frequently explain their decision making processes both to their entire class and to individual students. Reasons for "second chances" re-takes on quizzes, etc., are explained and justified as the teachers feel a need to do so.

IS CHARACTER DEVELOPMENT A PART OF THE ROLE OF AN ENGLISH TEACHER, AND IF SO, DOES THIS INFLUENCE GRADING, AND HOW?

All the teachers in this study asserted that character development is indeed part of being an English teacher. However, all insisted that they did not overtly judge the character of their students. Character building is subtle, and if it affects grades, the effect is equally subtle—or more so. **Nearly all the teachers admitted, for example, that if a student were "on the line" between passing and failing, their judgment of the child's character influenced their decision to add the few points or leave the child in the failing category.** While most saw this as an "end-of-the-marking-period" phenomenon, the highly subjective nature of many assignments in English class suggests that the same kind of decision can be made many times over the course of a marking period, so that the actual affect on grades may be greater.

The participating teachers reflected on the ways that character development might influence their grading, but clearly resisted the idea that they made broad judgments about the character of their students when it came to grading them.

DISCUSSION

Grading is heavily influenced by the values and beliefs of the teacher who grades (Buzzelli & Johnston, 2002). **Teachers' 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, 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 (Buzzelli & Johnston, 2002).

In addition to decisions about relative values of individual assignments and quality of work, other factors enter into grading decisions in high school English. Noddings, (1992, 2002) noted that 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, parents' wishes, and the satisfaction of one's own conscience, are all part and parcel of the everyday decision-making of teachers, and all of these issues are integrated into the process of determining grades.

Noddings (1995) argued 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) rested on the belief that the fundamental purpose of schooling is a moral one, and that "skills education" and "calls for excellence" in education failed 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.

In schools in the United States, the shaping of character has always been an implicit charge in the education of youth (McClellan, 1999). 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). The 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.

While courses for moral instruction can readily be found in religious schools (Jackson et al., 1993) the wide disagreements that exist across cultures and religions about what is moral means that the public schools, in spite of state mandates to train students in moral conduct, steer clear of direct moral instruction and carry out such instruction indirectly. 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. It is ironic that a matter of such seeming importance as morals is graded only implicitly at best.

LIMITATIONS

While Mellmax High School shares many similarities with other schools of its size and demographic profile, it is, like every school, unique, and thus is not representative of other schools in the scientific sense. As is true of any qualitative study, this study is not generalizable, but, insofar as Mellmax and its English teachers are similar to other schools, what is true of Mellmax High may be true of others. The purposive nature of the sample means that one may draw comparisons with other samples, but only to the extent that those samples are similar to this one.

A second limitation comes from the reliance on interviews as the chief means of data collection. Interviews allow informants to present themselves as they wish to be seen. There is no reason to doubt the veracity of the teachers' responses, but it is natural for people to present themselves in ways they believe are positive, and in that meet what they perceive as their listener's expectations. **The teachers may have actually believed they acted exactly as they said, but they may have only approximated their own desired conduct in actual practice.** The absence of observational data to support the interview data somewhat weakens this study.

A final limitation must also be acknowledged. Like all similar qualitative studies, this one required the sifting through, and classification of, enormous amounts of data. Like a lawyer preparing a case, a researcher must weigh out the relative importance of various bits and pieces of data and assemble them in a coherent and persuasive manner. Unlike a lawyer who sorts the evidence in order to favor the client, however, a researcher must guard against bias in sorting, classifying, and reporting the evidence. Despite

scrupulous effort to control bias, a researcher can never completely escape altogether from his or her own way of looking at things, and subtle preferences almost certainly color the researcher's choices of what to include and what to leave out. While every effort has been made to avoid this as far as possible in this study, the limitation remains.

FURTHER RESEARCH

Because this study focused on English teachers, further research might explore whether English teachers have different views of grading than teachers of other subjects and the degree to which the elements of the moral framework influence the grading practices of teachers of other subjects such as math and science.

While this study shows that teachers' expectations change and grading practices change with them, further research might look into ways in which teachers instructional practices are adjusted due to changing expectations—and how such changes in instruction then change evaluation practices and affect grading decisions.

A study exploring the extent to which the moral framework used here might affect grading in a post secondary setting (in a teacher preparation program, perhaps) would also be worth undertaking.

CONCLUSION

The supposition that “a grade is a grade” 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 assurance. This study began with assumptions that grading in an English classroom is a moral activity, that it can be a very involved process incorporating many finely focused and individualized considerations, and that the interpretation and accurate understanding of a grade requires an understanding both of the student receiving the grade and the teacher giving it.

Although 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.

The complex interconnectedness of the many elements of schooling ultimately finds some expression in the grades that appear on each

student's report card. Because of the complexity of moral issues in grading, a true picture of the teacher's demanding task of assigning grades on a regular, even daily, basis can only be accurately revealed by fully examining the moral issues involved and considering the whole picture rather than limiting its scope to things which are easily measured.

REFERENCES

- Baier, A. C. (1995). *Moral prejudices: Essays on ethics*. Cambridge, MA: Harvard University.
- Barber, T. X. (1969). Five attempts to replicate the experimenter bias effect. *Journal of Consulting and Clinical Psychology*, 33(1), 1-14.
- Bennett, W. J. (Ed.). (1993). *The book of virtues*. New York: Simon & Schuster.
- Beyer, L. (1997). The moral contours of teacher education. *Journal of Teacher Education*, 48(4), 245-254.
- Bracey, G. W. (1994). Grade inflation? *Phi Delta Kappan*, 76(4), 328-329.
- 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.
- 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.
- 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.
- Dewey, J. (1909). *Moral principles in education*. Carbondale, IL: Southern Illinois University Press.
- Dockery, E. R. (1995). Better grading practices. *Education Digest*, 60(5), 34-37.
- Ebel, R. L., & Frisbie, D. A. (1986). *Essentials of educational measurement (4th ed.)*. Englewood Cliffs, NJ: Prentice-Hall.
- Elashoff, J.D., & Snow, R.E. (1971). *Pygmalion reconsidered*. Worthington, OH: Jones.
- Feldman, A., Alibrandi, M., & Kropf, A. (1998). Grading with points: The determination of report card grades by high school science teachers. *School Science & Mathematics*, 98(3), 140-149.
- Gecan, C., & Mulholland-Glaze, B. (1993). The teacher's place in the formation of students' character. *Journal of Education*, 175(2), 45-57.
- Glasser, W. (1971). *The effect of failure on the life of a child*. Washington, DC: National Association of Elementary School Principals.
- Goodlad, J., Soder, R., & Sirotnik, K. (1990). *The moral dimensions of teaching*. San Francisco: Jossey-Bass.

- Hansen, D. T. (1997, March). Understanding students. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.
- Hargis, C. H. (1990). *Grades and grading practices: Obstacles to improving education and to helping at-risk students*. Springfield, IL: Charles C. Thomas.
- 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., Hopkins, B.R., & Stanley, J. C. (1990). *Educational and psychological measurement and evaluation (7th ed.)*. Englewood Cliffs, NJ: Prentice-Hall.
- Jackson, P.W., Boostrom, R. E., & Hansen, D. T. (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.
- Karmel, L.J. (1970). *Measurement and evaluation in the schools*. Toronto, ONT: Collier-MacMillan.
- Kerr, D. (1987). *Authority and responsibility in public schooling*. National Society for the Study of Education. Chicago: University of Chicago Press.
- Kirschenbaum, H., Napier, R.W., & Simon, S. B. (1971). *Wad-ja-get? The grading game in American education*. New York: Hart Publishing.
- McClellan, B.E. (1999). *Moral education in America: Schools and the shaping of character from colonial times to the present*. New York: Teachers College Press.
- Noddings, N. (1992). *The challenge to care in schools: An alternative approach to education*. New York: Teachers College Press.
- Noddings, N. (1995, January). A morally defensible mission for schools of the 21st century. *Phi Delta Kappan, 365-368*.
- Noddings, N. (2002). *Educating moral people: A caring alternative to character education*. New York: Teachers College Press.
- Rawls, J. (1971). *The theory of justice*. Cambridge, MA: Harvard University Press.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Self-fulfilling prophecies and teacher expectations*. New York: Holt, Rhinehart, and Winston.
- Sinclair, R. L., & Ghory, W. I. (1987). *Reaching marginal students: A primary concern for school renewal*. Berkeley, CA: McCutchan.
- Sockett, H. (1993). *The moral base for teacher professionalism*. New York: Teachers College Press.
- Thorndike, R. (1968). Review of Pygmalion in the classroom. *American Educational Research Journal, 5*, 708-711.
- Tom, A. (1984). *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.
- Viadero, D. (2003). Studies illuminate self-defeating behavior by students. *Education Week, 22*(28), 8.
- Wood, L. A. (1994). An unintended impact of one grading practice. *Urban Education, 29*(2), 188-202.

Copyright of American Secondary Education is the property of American Secondary Education and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.