

**FACTORS INFLUENCING GRADES AWARDED BY TEACHERS:
WHY DON'T GRADES REFLECT ACHIEVEMENT ONLY?**

by

Franklin A. Bruce, Jr.

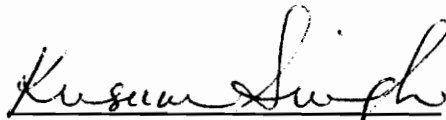
Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements
for the degree of

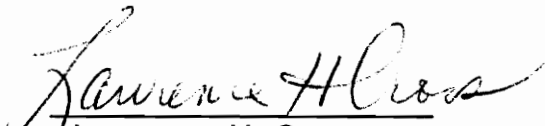
Doctor of Philosophy

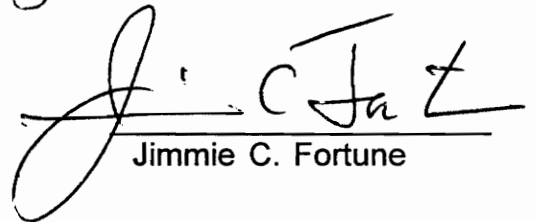
in

Educational Research and Evaluation

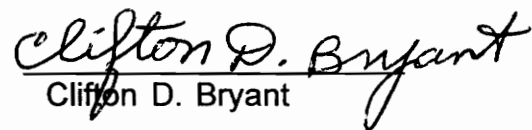
APPROVED:


Kusum Singh, Chairman


Lawrence H. Cross


Jimmie C. Fortune


Eileen S. Anderson


Clifton D. Bryant

January, 1995

Blacksburg, Virginia

C.2

LD

5055

V850

1995

5783

C.2

FACTORS INFLUENCING GRADES AWARDED BY TEACHERS: WHY DON'T GRADES REFLECT ACHIEVEMENT ONLY?

by

Franklin A. Bruce, Jr.

Committee Chairman: Kusum Singh
Educational Research and Evaluation

(ABSTRACT)

The purpose of the study was to identify the relative influence non-achievement factors have on secondary school teachers' grading practices. Specifically, an attempt was made to predict on the basis of non-achievement related factors the discrepancy between the grades teachers actually award to hypothetical student profiles and the grades that would result if based on academic achievement alone.

Teachers learn in their professional training to base grades on student achievement, however in actual practice they do not always follow such an approach. They openly acknowledge including non-achievement factors in awarding students grades, explaining that there are psychosocial consequences in grades that must be considered.

The study involved 192 secondary school teachers from twelve school districts who taught high school academic subjects. Teachers were asked to award grades to hypothetical student profiles in simulated grading situations. Each simulation, or vignette, presented a hypothetical student profile and included a unique combination of factors which had been shown in previous

research to influence teachers' grading decisions. A technique referred to as policy capturing was used to evaluate the extent to which assigned grades reflect student gender, effort, behavior, ability, and test score improvement.

In predicting the grade discrepancies of teachers, it was judged necessary to restrict the number of grade discrepancies for teachers to five or more, which represented approximately one-sixth of the possible 32 discrepancies provided by the vignettes, reducing the number of teachers in the sample to 99.

Regression models predicting the 99 teachers' grade discrepancies, which were differences between the letter grade awarded by teachers and the grade implied by the test scores in the vignettes, and policy capturing analysis revealed that student test score improvement and student behavior influenced teachers most in their grading. However, teachers seemed to place more importance on students' quiz average grades as a measure of academic performance, than on students' six-week test grades. In addition, secondary analyses of the data revealed that in predicting the importance placed by the 99 teachers on students' test score improvement and on behavior, teachers with more years of experience placed less importance on either of the factors in grading than their less experienced colleagues.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. Kusum Singh, Dr. Lawrence Cross, Dr. Jimmie Fortune, Dr. Eileen Anderson, and Dr. Clifton Bryant for their support and assistance in completing this study. I am particularly appreciative to Dr. Lawrence Cross, who provided me with encouragement when I needed it, and constructive criticism when it was due. Also, I would like to thank Dr. Kusum Singh for chairing this committee and for her patience, encouragement, and belief in me as a researcher. I would also like to acknowledge Dr. Bob Frary for his assistance and advice in the early stages of writing the proposal and in reviewing early drafts.

I would like to acknowledge all the superintendents and administrators in the school districts who graciously permitted and assisted me in inviting teachers to participate in the study, and to all the teachers who took time from their schedules to participate in the exercise.

Finally, special thanks to my family, Janet and Jake, in providing inspiration and encouragement during difficult times, and to my parents for their everlasting patience, encouragement, and confidence in me.

TABLE OF CONTENTS

Abstract	ii
Acknowledgements	iv
List of Tables	vii
List of Figures	viii
Introduction	1
Review of Selected Literature	3
Purpose of the Study	10
Methodology	11
Subjects	11
Policy Capturing	12
Instrument	15
Pilot Testing	15
Informational Elements and Their Levels	16
Procedure	17
Analyses	19
Data Analysis	20
Weights	21
Teachers' Characteristics	22
Results and Discussion	24
Teachers with Five or More Grade Discrepancies	25
Most Frequently Occurring Significant Factors	33
Policy Capturing Weights	36
Teachers' Characteristics	37
Frequencies of Grade Discrepancies by Vignettes	41
Teachers' Comments	46
Conclusions	53
Effect of Grade Change	56
Effect of Student Behavior	57
Implications and Recommendations for Future Research	60
References	65
Appendix	
A. Time Line of Data Collection	69

B.	Factor Levels and Coding	71
C.	Review of Selected Literature	89
D.	Teacher Comments	105
Vita		127

List of Tables

<u>Table</u>		<u>Page</u>
1	Number of teachers, average years of teaching experience, educational level, and gender of teachers responding to exercise on grading practices by subject area	13
2	Frequencies of teachers' awarded grades by vignette	26
3	Number of teachers by number of grade discrepancies for the 32 vignettes	27
4	Number of teachers responding to exercise on grading practices by subject area with five or more grade discrepancies	28
5	Influence of non-achievement factors on secondary teacher's grading practices: standardized regression coefficients for each teacher	29
6	Number of teachers for whom a student informational factor was a significant predictor of discrepancies between grades awarded and grades based on academic performance	34
7	Teachers' characteristics that predict importance placed on the non-achievement factors: student behavior and grade change	42
8	Number of discrepancies, percentages, corresponding vignette number, and profile coding of the most positive and negative grade discrepancies by teachers	43
B-1	Vignette factor level combinations, letter grade changes, and numeric and letter grade averages by grade change	86
B-2	Presentation Order for Vignettes in Forms 1, 2, and 3	87
C-1	Summary of policy capturing literature in terms of study's characteristics	103

List of Figures

<u>Figure</u>		<u>Page</u>
1	Hypothetical student profile	18
2	Policy capturing weights of teachers having five or more grade discrepancies	38
3	Content and Organization of Vignettes Receiving Highest Number of Positive and Negative Grade Discrepancies from Teachers (N = 192)	47
B-1	Vignettes and Factor Level Combinations	75
B-2	Directions and Questions Attached to Each Packet of Vignettes	88

Introduction

In their professional training, teachers "dutifully learn the recommended practice of basing grades on achievement" (Brookhart, 1991, p.36), rather than on non-achievement factors such as student effort, ability, or attitude. However, in actual practice teachers do not always follow such a strict approach, at least at the secondary level (Frary, Cross, & Weber, 1993; Nava & Loyd, 1992; Stiggins, Frisbie, & Griswold, 1989). Teachers point out that basing grades solely on achievement "won't always work" and that, sometimes, non-achievement factors are also considered (Brookhart, 1991).

Why do teachers feel they must include noncognitive factors in grading? Brookhart (1991) suggests that it is because "teachers know that grades are not only interpreted, they are used" (p.35). That is, while teachers have been taught that their assigned grades should represent what a student has learned, many say they use assigning grades "for lots of different things" (p.36). For example, since teachers are aware that grades can have both desirable and undesirable social consequences, they may alter a student's semester or six week grade from "D" to "C" allowing that student to participate in a certain extracurricular activity, or assign another student a "D" rather than an "F" in an attempt to influence that student not to drop out of school (Brookhart, 1991).

Furthermore, grades represent an important communication link between teachers and other persons, e.g., students, parents, counselors,

administrators, colleges, employers, etc. Assigning grades is intended "to provide clear, interpretable feedback... about...academic learning (Brookhart, 1991, p.39). If grades represent something other than a student's academic performance, such as effort, the act of factoring it "into an academic grade equation may have the simultaneous effects of (a) encouraging motivation, and (b) inhibiting clear communication between the grader who sends the message via receiver down the road who must act on the information provided in the grade" (Stiggins, Frisbie, & Griswold, 1989, p.14).

If the basis on which a grade is awarded is not clear, there is room for ambiguity and misinterpretation. Take for example a student receiving a grade of "B" for a semester's work in algebra. Does the grade mean the student knows a lot about algebra, or that the student is a hard worker? If grades are based on factors other than achievement their validity, what they mean, is obscured. "One does not average oranges and bicycles; it is wise to keep marks for achievement separate from ratings of study habits and attitudes" (Hopkins, Stanley, & Hopkins, 1990, p.328).

While a system operating on a "hodgepodge grade of attitude, effort, and achievement created in an attempt to provide positive feedback...is not the answer" (Brookhart, 1991, p.36), it does seem to happen and may be happening a lot (Frary et al., 1993; Nova & Loyd, 1992). Airasian (1991) points out that "there is pressure on the teacher not only to provide an accurate report of pupil performance, but also to minimize the negative

repercussions of grades" (p.309). He believes that the reason teachers have mixed feelings on the subject is that teachers are aware of: (1) the broad range of uses found for grades (e.g., class rank, credits for graduation, grade promotion, informing parents, etc.); (2) grades representing a teacher's summary of learning in a subject and also behaviors taught in that subject; (3) grades alone not providing detailed diagnostic information about a student's performance; (4) use of grades to motivate students to study; and (5) use of grades for assisting students, parents, and counselors in choosing appropriate courses and course levels for students.

Measurement specialists may cringe at the "hodgepodge" of factors considered by some teachers in assigning grades. However, in reality many teachers say they vary the importance they assign non-achievement grading criteria and say they do so at different times for different reasons (Frary et al, 1993; Nova & Loyd, 1992). While research tells us such practice exists, identifying what influence non-achievement related factors have on the grading process needs to be examined.

Review of Selected Literature

The research on teachers' use of non-achievement factors in assigning grades is rather limited but increasing. A recent review of teachers' grading practices was conducted by Brookhart (1994). She reviewed findings of 19 studies conducted during the last 10 years involving grading practices of

classroom teachers (K-12) who taught academic subject(s) in basic education. Findings from the studies include teachers: (1) attempting to be fair to students, informing them of what the components of their grades will be before awarding grades, (2) considering student effort and ability along with achievement measures in grading, (3) differing, according to whether they teach elementary or secondary school students, in what the components of a grade may be, and (4) considering achievement and nonachievement factors differently in grading. Furthermore, she reported that her findings indicated a "lack of congruence between recommendations of measurement specialists and classroom teachers' grading practices" (p.289).

In studying teachers' actual grading practices, Barnes (1985) examined attitudes, knowledge bases, and practices of twenty supervisory teachers and twenty student teachers who were placed with them regarding the pupil process evaluation. Both supervising and student teachers said "grading and evaluation were the most difficult aspects of teaching" (p.47). The study found all the teachers struggling with "a pervasive and unresolved conflict....between evaluation on the basis of performance and evaluation on the basis of effort" (p.48). Typical comments reported by supervising teachers were "A lot of times I grade on effort rather than outcome", or "it's a real conflict...I have to put number grades for averages but I try to emphasize...the reason we're doing this is so we can check and fix it and know better next time" (p.48). While the study offers "some valuable insights"

of teachers, it and other studies addressing non-achievement factors in grading have been limited by such things as small sample sizes, lack of random selection of subjects, and combining vocational and academic subject teachers (Frary et al., 1993, p.23).

Agnew (1985) examined grading policies and practices of two hundred and sixty-six high school teachers from seven different schools. The study found two non-achievement factors, the amount of improvement a student had shown and how hard a student worked, considered by teachers as important to take into consideration when assigning grades. The study asked teachers how important "certain referents were in the assignment of (their) course grades" (p.22). Referents included: (1) "how a student's work compared with other students' work in the same class (norm-referenced)", (2) "whether individual student work met or meets the grade level criteria set for all students (criterion referenced)", (3) "the amount of improvement the student has shown", (4) "how hard the student worked", and (5) "how well the student behaved in class" (p.22). Teacher's opinions were solicited using a 5-point "not important/extremely important" response scale. The study found the most important referent to be "meets grade level criteria (criterion referenced)" (mean = 3.87, SD = 1.07); however, the second and third most important referents were "the amount of improvement the student had shown" (mean = 3.55, SD = 1.01) and "how hard a student worked" (mean = 3.30, SD = 1.1), respectively. Of the remaining two, "comparison of the student's

work with others in class (norm-referenced)" (mean = 2.82, SD = 1.23) ranked fourth and student behavior ranked last (mean = 2.61, SD = 1.13).

A study by Manke and Loyd (1990) found, when teachers were asked to make "grading decisions" from hypothetical student simulations, teachers used the factors student effort, academic improvement, behavior, personality, homework, and gender, all non-achievement factors, in making decisions. The study included one hundred and six teachers (elementary, 53; secondary, 53) as subjects. Importance was attributed to a non-achievement factor by the percentage of teachers who used the factor when it was presented as an option in determining a hypothetical student's grade. Each of seven non-achievement factors was presented in a separate "problematic grading situation" simulation. In order of importance, teachers considered, first, a student's completion of homework (97%), second, if a student exhibited high effort in class (89%), third, a student's personality (88%), fourth, a student's behavior (78%), and fifth, if a student exhibited low effort (48%). In addition, differences were found between elementary and secondary teachers in the importance each placed on the factors: high effort, student personality, grade improvement, and student interest.

The study asked teachers to read eleven hypothetical simulations and for each make a "grading decision". As an example, one simulation depicted a student with average ability and not seeming to be very interested in her schoolwork. The statement "You would like to motivate her to try harder" and

a grade average of "D" was also included. Responses teachers could choose included: (1) "lower the student's grade", (2) "raise the student's grade", or (3) "grade the student on achievement alone". Sets of response choices were different for each of the simulations, ranging from two to five choices. Non-achievement factors were used in only seven of the eleven simulations, the other four dealt with conflicts between two aspects of achievement.

The Manke and Loyd study also found a student's gender to influence teachers judgement when assigning grades. The authors created an alternate form of the eleven simulations by reversing the gender of the students in the seven non-achievement simulations. In two simulations there was a significant relationship between teachers responses and student gender. The two simulations were: "a student who showed high achievement on tests but did not do homework was more likely to be penalized if the student was male; and a student who showed substantial improvement on tests but remained within the F range was more likely to have improvement rewarded with a grade of D if the student was female" (p.30). The authors reported "that at least in some situations the gender of the student determines what the teacher believes is appropriate for the grade" (p.30).

A study investigating the importance teachers place on achievement and non-achievement grading criteria was conducted by Nova and Loyd (1992). They surveyed eight hundred and twenty seven teachers (elementary, 371; secondary, 456) from 18 school districts from various

regions of the country. Teachers were asked to rate thirty-five achievement and non-achievement grading criteria "according to the degree that each one should be included in grading" (p.4). Ratings were on a 4-point definitely include/definitely not include response scale. The study found teachers responding to twenty-three of the thirty-five criteria as either "definitely include in grading" or "probably include in grading". The two most responded to criteria were unit tests and announced quizzes (94.7% and 91%, respectively); however, the third most responded to criterion item was student effort (88.3%). These criteria were considered by the authors as "the most important criteria that teachers report they 'should include' in grading" (p.6). Of the remaining twenty criteria, twelve included non-achievement measures. Included were improvement during the grading period (83.1%), completing work on time (82.6%), homework/assignments (81.8%), participating in class (78.1%), improvement from one grading period to the next (74%), ability level (73.1%), participating in group discussions (73%), cheating (70.9%), cutting/skipping class (55.9%), cooperativeness (54.5%), regular attendance (53.4), and inattention in class (50.2).

The study concluded that most teachers used a variety of achievement and non-achievement factors in assigning grades and varied the importance of each. In addition, the study analyzed teachers' ratings of criteria using exploratory principal components factor analyses. A four factor analysis design was chosen. The four factors were reported as: (1) "achievement or

classroom content-related item (e.g., tests, quizzes, projects, papers); (2) criteria that are either characteristics or behaviors that facilitate or interfere with learning (such as effort, cooperativeness, participation, ability level, inattention); (3) characteristics that are academic in nature but not necessarily related to specific classroom content (such as spelling, grammar, handwriting, attendance, cutting/skipping class, completing work on time); and (4) characteristics that are related to overall school or classroom performance but are external to the classroom (such as parental involvement in class and school activities, gender and socio-economic status of the student" (p.19).

Frary, Cross, and Weber (1993) conducted a study which elicited secondary school teachers' views on testing and grading practices. In the spring of 1991, all secondary teachers of academic subjects employed in Virginia public schools were randomly sampled. The sample consisted of five hundred and thirty-six teachers representing a sixty-seven percent response rate. Included were only teachers having at least a fifty percent teaching assignment in any combination of English, mathematics, science, social studies, or foreign languages. The authors constructed a 44-item questionnaire which included seventeen items asking "factual information" (about testing and grading practices) with the remaining items "soliciting opinions about testing and grading in an academic course" (p.24). The study found teachers reporting, as "important" factors influencing grading, a student's daily homework and class participation (71% and 51%,

respectively). When teachers were asked "their opinion" concerning other non-achievement factors, they felt a student's knowledge gains over an instructional period (85%), ability (66%), exceptional low or high effort (66%), and laudatory or disruptive classroom behavior (35%) should be taken into consideration in awarding grades.

As Nava and Loyd (1992) point out, grading for teachers "has become a complex, varied, subjective procedure ... little is known about the specific criteria that teachers, in general, include in grading, or how the specific criteria are evaluated and used in combination when teachers make decisions on a student's end-of-the-term grade" (p.21).

Purpose of the Study

The purpose of this study was to identify the relative influence non-achievement factors had on secondary teachers' grading practices. Specifically, an attempt was made to predict on the basis of non-achievement related factors the discrepancy between the grades teachers actually award to hypothetical student profiles and the grades that would result if based only on academic achievement.

Methodology

Subjects

The study included 192 secondary school teachers who taught at least one academic class. Twelve school divisions within the state of Virginia agreed to allow teachers from all or some of their high schools to be invited to participate in the study. In four school districts exercise packets were either mailed or hand delivered to an administrator who distributed the packets among high school teachers they thought would be interested in the study. In two districts, packets were placed in the mailboxes of teachers school principals felt would be interested in the study. In the remaining five school districts, teachers were presented the exercise by the researcher either while in their classrooms or when they attended a teachers' meeting.

In a letter of introduction attached to each of the exercise packets, teachers were informed that there was no obligation for them to participate in the exercise, but that doing so might stimulate them to think more specifically about their personal grading practices. Participating teachers were from small, medium, and large county schools as well as from two city schools. The time line for data collection, participating school districts/schools, manner in which teachers were contacted, and number of returned exercises is provided in Appendix A.

Teachers were asked to respond to 32 hypothetical vignettes with regard to one of their academic classes. Eight different academic subjects were reported by teachers. The number of teachers reporting each of the subject areas and descriptive information about the teachers is presented in Table 1. Of the 192 teachers, seventy-two (38%) were male and 117 (62%) were female. Three of the teachers did not record their gender. One hundred and eleven teachers (58%) had either a BS or BA degree, seventy-nine (41%) had earned a masters degree, and two had earned doctoral degrees. Teachers ranged in teaching experience from 1 to 36 years (Mean = 17.21, SD = 9.48).

It was not the intention of the study to include solely white teachers in the sample. However, one hundred and eighty-two (95%) of the 192 teachers were white. Nine teachers (5%) reported their ethnicity as other than white. Of these, two were African-American, one was an Asian/Pacific Islander, two were Hispanic, one was American Indian/Native Alaskan, and three marked the "Other" category. One teacher did not record his/her ethnicity.

Policy Capturing

The questions addressed in this research can be viewed within a decision-making context. Put simply, teachers sometimes decide to assign a particular student a grade based on both achievement and non-achievement related information they have regarding the student. Studies addressing

Table 1

Number of Teachers, Average Years of Teaching Experience, Educational Level, and Gender of Teachers Responding To Exercise On Grading Practices By Subject Area

(N=192)

<u>Subject Area</u>	<u>Number of Teachers</u>	<u>Average Years Experience</u>	<u>Educational Level</u>	<u>Gender</u>
Math	49	17.79 (SD=9.24)	61.2% BS 38.8% Masters	40.8% Male 59.2% Female
English	49	18 (SD=9.47)	59.2% BS 40.8% Masters	14.9% Male 85.1% Female
Science	40	16.85 (SD=10.12)	40% BS 60% Masters	50% Male 50% Female
History	21	19.35 (SD=8.474)	70% BS 30% Masters	65% Male 35% Female
Government	10	14 (SD=8.817)	66.7% BS 33.3% Masters	66.7% Male 33.3% Masters
Foreign Language	19	12.578 (SD=9.985)	63.2% BS 36.9% Masters	21.1% Male 78.9% Female
Social Studies	3	23.333 (SD=4.041)	66.7% BS 33.3% Masters	33.3% Male 66.7% Female
Psychology	1	21	100% BS	100% Male
Total	192			

teacher decision-making have used a research methodology approach called "policy capturing" (e.g., Borko & Caldwell, 1982; Tomchin, 1989). Hobson and Gibson (1983) characterize the approach as "a general procedure designed to describe statistically the unique information processing strategies of individual raters" (p.640). They describe the "policy capturing" approach as consisting of the following steps: "(a) the presentation to raters of a series of performance profiles consisting of scores on the major dimensions of performance; (b) instructions to raters to review each profile and then assign an overall rating that best represents or summarizes the available information; and (c) the use of multiple regression analysis to calculate the extent to which each individual rater's overall ratings are predictable, given scores on the separate performance dimensions, and to compute the relative importance of each single dimension in determining overall ratings" (p.640).

In using "policy capturing" methodology in this study, the "performance profiles" noted above were hypothetical student grading situations. Each grading situation described a different student with a unique set of characteristics, behavior, and academic performance. Teachers selected for the study were asked to read each of the hypothetical grading situations and assign a grade to the student depicted in each situation.

In reviewing policy capturing research, Jaeger and Usher (1991) found simulation to be one of the most popularly used methods for developing performance profiles (See e.g., Hobson, Mendel, & Gibson, 1981; Stumpf &

London, 1981; Tomchin, 1989). Tomchin (1989) explained how simulated student profiles were created for her study involving teachers' grade retention policies, "a number of cues, in this case characteristics of students, are varied systematically in vignettes of hypothetical students. The retention decisions made about each student are analyzed to determine the importance of each cue or combination of cues presented in the series of vignettes" (p.32). Additional information concerning policy capturing methodology is provided in Appendix C.

Instrument

The major dimensions of performance, or informational elements, in this research were characteristics, behaviors, and performance of students. These were all factors which had been shown in previous research as student characteristics teachers take into consideration when assigning grades. Factors or informational elements hypothesized to be considered by the teachers were student gender, effort, behavior, ability, and changes in two academic achievement grades.

Pilot Testing

Thirty-two simulated hypothetical grading situations, called vignettes, were created along with directions and pilot tested by 12 high school teachers who taught academic subjects. There were four teachers selected from each

of three nearby high schools. The 32 vignettes were divided into blocks of eight, each teacher in a school receiving eight different vignettes. Teachers were asked to read the directions and the vignettes they received, award grades to the students depicted, and offer comments concerning the "believability" of the various factor level and grading scenarios. Each teacher in a school received only eight vignettes for several reasons. These included: so comparisons could not be made by teachers in a school regarding grades awarded to vignettes, to promote closer inspection of the vignettes, and to initiate comments from teachers.

Teachers' comments included that the students depicted in some of the vignettes were "just like" some of their own students. Also, the majority of the teachers participating in the pilot study were found awarding grades which were different than the grades implied by the test scores in the vignettes.

Informational Elements and Their Levels

A vignette described a hypothetical student as either male or female, exhibiting either low or high effort, behaving in an unacceptable or desirable manner in class, and with either low or high ability. In addition, each hypothetical profile depicted a student having acquired two academic achievement grades for a six-week grading period, a quiz average grade (average of four quizzes taken during the six-weeks) and a six-week test grade. These two grades were varied to include increases and decreases

from the quiz average grade to the six-week test grade. Informational elements and their various levels varied within vignettes were:

1. Sex (1=male, 2=female),
2. Effort (1=low, 2=high),
3. Behavior (1=unacceptable, 2=desirable),
4. Ability (1=low, 2=high),
5. Test score improvement (grade change) - grade change from quiz average grade to six-week test grade (1=increase, 2=decrease).

Thirty-two hypothetical student profiles were constructed to satisfy all possible combinations of the above two levels for each of the five informational elements. Construction was modeled after other research using vignettes (Borko, 1978; Tomchin, 1989). A sample vignette is shown in Figure 1.

It was also decided to create three different forms of the exercise, or ordering of vignette presentation, to guard against teachers making comparisons of each others' responses. The order of the three vignette presentations is provided in Appendix B. See also, Appendix B for how factor terms were interpreted, the thirty-two vignettes created, and the directions each teacher received.

Procedure

First, teachers were asked to read the directions for the exercise and then complete a brief questionnaire which gathered information regarding the number of years they had taught, educational attainment, subjects taught, and

A Sample Vignette

Robert never wants to do any more than he absolutely has to in your class. He has gym just before your class and often has to hurry to get there on time, but he always makes it. However, he seldom has his homework done, always wanting more time to complete it. He offers excuses such as he had to go visit a sick relative or left his book at school, etc. He had a D+ average on the four quizzes you gave during the six-weeks and a C on the six-week test. You think he's much smarter than his grades indicate. Other teachers have told you that he does very well in their classes. What grade would you assign Robert for the six-weeks? A, B, C, D, F

Figure 1. Hypothetical Student Profile

grade levels. The questionnaire is provided along with the directions in Appendix B.

Second, each teacher was asked to respond to each of the thirty-two vignettes by awarding the student depicted in each a grade. Therefore, the hypothetical student in each vignette received a letter grade (A, B, C, D, or F) from each of the participating teachers.

The directions explained to teachers to think of themselves as teaching one of their academic subject areas in responding to the vignettes. In the case of teachers who taught more than one subject area or different grade levels, they were asked before responding to the vignettes to pick a specific subject area and grade level and record the information accordingly.

Analyses

The grade each teacher assigned a hypothetical student profile was assigned a value relative to its proximity to the grade the student in the profile would receive if the grade were based only on achievement. If the two grades were the same the value assigned was zero. If a discrepancy existed between the assigned grade and achievement only grade, a +1, +2, +3, or +4 was recorded according to whether the assigned grade was 1, 2, 3, or 4 letter grades higher than the grade based on achievement only. Conversely, a -1, -2, -3, or -4 was recorded according to whether the assigned grade was 1, 2, 3, or 4 letter grades lower than the grade based on achievement only. For

example, if a teacher assigned a grade of "D" to a hypothetical student in a vignette and the numerical average of that student's quiz average grade and test grade was "B", the teacher would receive a discrepancy score of -2 for that vignette since the teacher's assigned grade was lower than the achievement only grade by two letter grades.

In this way each teacher received a "discrepancy score" for each of the thirty-two vignettes. As stated earlier, a teacher could receive a "discrepancy score" of zero, which means there was not any discrepancy between the grade they assigned and the student's achievement only grade.

Data Analysis

A separate regression equation was calculated for the 99 teachers with five or more grade discrepancies regressing each teacher's grade discrepancy scores from the thirty-two vignettes (N=32) on the five factors (sex, effort, behavior, ability, and test score improvement). Jaeger and Usher (1991) point out that "multiple linear regression is the analytic procedure of choice throughout the policy capturing literature" (p.14). In addition, Slovic and Lichtenstein (1971), in their review of policy capturing and comparable methodologies, found that in policy capturing studies regression models incorporating interaction terms showed "extremely small increments in proportions of predicted variation of criterion scores compared to models that are strictly additive" (Jaeger & Usher, 1991, p.14). In his review of the policy

capturing approach, Jaeger and Usher (1991) cite research supporting the "robustness of additive linear models in the behavioral decision-making theory literature" (p.14) justifying the elimination of interaction terms. Therefore, the regression model used in this study assumed that the informational elements (factors) provided teachers contributed linearly and additively to discrepancies in the grades teachers assigned and grades based on achievement only.

Weights

In conducting a policy capturing study, Jaeger and Usher (1991) calculated regression equations for each individual rater in their study and then calculated attribute "relative weights" for each regression coefficient in each rater's equation (Hoffman, 1960). Calculation was as follows:

$$w_i = B_i r_i / R^2$$

w_i = weight for the i th attribute

B_i = standardized regression coefficient

r_i = correlation between the dependent variable and the i th attribute

The advantage of calculating "relative weights" is that they can be compared both across teachers and across factors (informational elements). In presenting the results of their study, Jaeger and Usher made such comparisons. Median values were calculated for each distribution of weights for each attribute. These weights, representing responses of raters to the hypothetical profiles, were then compared across raters and across attributes (factors) in box-and-whisker charts. In this study, after a separate

regression equation was calculated for each teacher, the regression coefficients in each teacher's equation were converted to a "relative weight," using Hoffman's formula. To examine the importance placed by teachers on the various informational elements, distributions and median values of the weights for each informational element was calculated and presented in box-and-whisker charts.

Several policy capturing studies have examined only the significant ($p < .05$) standardized regression coefficients resulting from each subject's separate regression equation for systematic variation (e.g., Borko & Caldwell; Taylor & Wilsted, 1974; Tomchin, 1989) rather than assign weights as Hoffman (1960) suggests. In the present study, both standardized regression coefficients and Hoffman's weights were calculated.

Teachers' Characteristics

Teacher characteristics (e.g., years of teaching experience, gender, subject area) were examined by using information obtained about each of the teachers. For teachers with five or more grade discrepancies (N=99), a series of multiple regression analyses were performed using the regression weights for each of the informational elements from each teacher's regression equation as dependant variables and teacher characteristics as independent variables (See e.g., Borko & Caldwell, 1982). Borko and Caldwell (1982) used such analyses in a study concerning teachers' classroom management

practices. They reasoned that any teacher characteristics having an impact on a teacher's decision policy would be statistically significant in one or more of the equations predicting the various regression weights.

Results and Discussion

The criterion variable of interest for the study was computed as the difference between the letter grade awarded and the grade implied by the test scores in the vignettes. These discrepancies were coded as either -4, -3, -2, -1, 0, +1, +2, +3, or +4 corresponding to whether the awarded grade was one, two, three, or four letter grades above (positive) or below (negative) the student's academic performance average.

Within the 32 vignettes, five student informational elements or factors were systematically varied. Factors, along with their respective levels, include: (1) gender (1 = male, 2 = female), (2) effort (1 = low, 2 = high), (3) behavior (1 = unacceptable, 2 = desirable), (4) ability (1 = low, 2 = high), and (5) grade change from the student's quiz average grade to their six-week's test grade (1 = decrease, 2 = increase).

In awarding grades to the 32 vignettes, the 192 teachers recorded many more grade discrepancies for some vignettes than others. For example, three vignettes (#2, #18, and #20) received grade discrepancy totals from teachers of 6, 3, and 7, respectively, while three other vignettes (#13, #7, and #19) received discrepancy totals of 105, 98, and 96 discrepancies, respectively. For most vignettes teachers' grade discrepancies tended to be either mostly higher or lower than the grade implied for a vignette. However, in four vignettes (#1, #4, #18, and #30), where there were few grade discrepancies, there were approximately equal numbers of discrepancies

above and below the implied grades. A frequency distribution of teachers' awarded grades to each of the 32 vignettes is provided in Table 2.

In addition, the majority of the 192 teachers (93.2%) had at least one grade discrepancy. Only ten teachers (6.8%) did not have any grade discrepancies. The number of teachers by grade discrepancy is provided in Table 3.

Teachers With Five or More Grade Discrepancies

A regression equation was created for each teacher who had five or more grade discrepancies across the 32 vignettes. The decision to include only teachers having five or more discrepancies was a judgement call that recognized the need for variance on the criterion variable. Clearly, if there were no grade discrepancies for a teacher, there would not be any criterion variance to be predicted. To require more than five discrepancies would reduce the number of available cases severely. Ninety-nine teachers were found to have at least five grade discrepancies, representing 51.6% of the 192 participating teachers. The number of teachers with five or more grade discrepancies by subject area is provided in Table 4.

In order to show the relative importance each of the five factors played in the grades awarded by a teacher, standardized regression coefficients were interpreted. The five standardized regression coefficients for each of the 99 teachers are shown in Table 5. Those factors which were significant

Table 2

Frequencies of Teachers' Awarded Grades by Vignette

(N=192)

Vignette/Levels	Implied Grade	A	B	C	D	F	Total Number of Discrepancies
#1 (1,1,1,1,1)	D	1	0	8	169*	14	23
#2 (1,1,1,1,2)	D	1	0	5	186*	0	6
#3 (1,1,1,2,1)	C	1	9	176*	6	0	16
#4 (1,1,1,2,2)	C	0	2	185*	5	0	7
#5 (1,1,2,1,1)	D	0	0	15	173*	4	19
#6 (1,1,2,1,2)	B	1	165*	26	0	0	27
#7 (1,1,2,2,1)	D	1	0	97	94*	0	98
#8 (1,1,2,2,2)	C	0	0	136*	55	1	56
#9 (1,2,1,1,1)	C	0	2	175*	15	0	17
#10 (1,2,1,1,2)	D	0	0	2	158*	32	34
#11 (1,2,1,2,1)	C	0	7	175*	10	0	17
#12 (1,2,1,2,2)	B	0	171*	21	0	0	21
#13 (1,2,2,1,1)	C	1	104	87*	0	0	105
#14 (1,2,2,1,2)	B	4	170*	18	0	0	22
#15 (1,2,2,2,1)	C	1	50	140*	1	0	52
#16 (1,2,2,2,2)	B	43	149*	0	0	0	43
#17 (2,1,1,1,1)	D	0	0	11	178*	3	14
#18 (2,1,1,1,2)	C	1	1	189*	1	0	3
#19 (2,1,1,2,1)	D	0	2	90	96*	4	96
#20 (2,1,1,2,2)	D	1	0	2	185*	4	7
#21 (2,1,2,1,1)	C	0	1	181*	10	0	11
#22 (2,1,2,1,2)	C	0	2	171*	19	0	21
#23 (2,1,2,2,1)	C	0	1	172*	19	0	20
#24 (2,1,2,2,2)	B	20	171*	1	0	0	21
#25 (2,2,1,1,1)	B	3	173*	16	0	0	19
#26 (2,2,1,1,2)	B	0	160*	31	1	0	32
#27 (2,2,1,2,1)	C	0	24	162*	6	0	30
#28 (2,2,1,2,2)	B	0	142*	47	0	3	50
#29 (2,2,2,1,1)	C	1	56	134*	1	0	58
#30 (2,2,2,1,2)	C	0	8	178*	6	0	14
#31 (2,2,2,2,1)	B	3	182*	7	0	0	10
#32 (2,2,2,2,2)	B	30	161*	1	0	0	31

Note. Factor order and coding of vignettes are: sex (male=1, female=2), effort (low=1, high=2), behavior (unacceptable=1, desirable=2), ability (low=1, high=2), and grade change (decrease=1, increase=2).

* Frequency of no grade discrepancies for the vignette

Table 3

Number of Teachers by Number of Grade Discrepancies for the 32 Vignettes

(N=192)

<u>Number of Grade Discrepancies</u>	<u>Number of Teachers/Percentage</u>
0	13 (6.8%)
1	21 (10.9%)
2	21 (10.9%)
3	22 (11.5%)
4	16 (8.3%)
5	26 (13.5%)
6	15 (7.8%)
7	12 (6.3%)
8	8 (4.2%)
9	12 (6.3%)
10	7 (3.6%)
11	4 (2.1%)
12	4 (2.1%)
13	4 (2.1%)
14	none
15	1 (.5%)
16	1 (.5%)
17	3 (1.6%)
18	2 (1%)
	<hr/>
Total	192

Table 4

Number of Teachers Responding To Exercise On Grading Practices By Subject Area With Five or More Grade Discrepancies

<u>Subject Area</u>	<u>Number of Teachers With Five or More Grade Discrepancies</u>
Math	26
English	28
Science	20
History	11
Government	5
Foreign Language	2
	<hr/>
Total	99

Table 5

Influence of Non-achievement Factors on Secondary Teacher's Grading Practices: Standardized Regression Coefficients for Each Teacher (N=99)

<u>Teacher</u>	<u>Student Informational Elements</u>					<u>Number of Discrepancies</u>
	Gender	Effort	Behavior	Ability	Gr/Change	
1	.000	-.144	.144	.000	-.433*	6
3	-.086	.086	.258	.086	-.258	5
4	.079	-.079	.158	.317*	-.634*	17
5	.079	.079	-.079	.079	-.396*	5
6	.059	-.059	.059	.177	-.531*	9
7	.250	.000	-.125	.250	-.500*	8
8	-.061	.061	.308	.061	.061	9
9	.244	-.081	-.081	-.081	.081	5
10	.049	-.049	-.049	.245	-.638*	13
11	.088	-.177	.177	.266	-.532*	13
15	.000	.437*	.145	.000	-.145	6
18	.000	-.086	-.086	.086	-.774*	18
22	.000	.160	.480*	.160	.000	6
23	.000	.160	.160	.160	-.480*	6
25	.081	.081	-.244	-.081	-.244	5
28	.129	.129	.258	.000	-.129	12
30	.102	.204	.306	.102	-.306	12
32	.067	-.067	.200	-.200	-.468*	7
37	-.226	-.226	-.075	.075	.075	7
40	-.160	.320*	.480*	.160	-.160	6
43	-.086	.086	.258	.258	-.086	5
44	.000	.000	.443*	.443*	.088	16
45	-.179	-.059	.059	.059	-.419*	9
46	.151	-.050	.050	.050	-.455*	9
47	.244	.081	.244	.081	-.244	5
48	-.043	.562*	.302*	.129	.129	17

Note. Regression equations were calculated for teachers who had five or more grade discrepancies.

Teachers' grade discrepancies were coded -4, -3, -2, -1, 0, +1, +2, +3, or +4 according to whether a teacher's assigned grade was one, two, three, or four letter grades below or above the average of a student's academic performance.

Teacher's grade discrepancies were regressed on Student Informational Elements.

Table 5 (Cont'd.)

Influence of Non-achievement Factors on Secondary Teacher's Grading Practices: Standardized Regression Coefficients for Each Teacher (N=99)

<u>Teacher</u>	<u>Student Informational Elements</u>					<u>Number of Discrepancies</u>
	Gender	Effort	Behavior	Ability	Gr/Change	
51	-.129	.129	.258	.258	-.129	8
55	-.129	.129	.258	.000	-.387*	8
60	-.295	-.059	-.177	-.295	-.295	9
62	-.053	-.053	-.053	.160	-.587*	11
66	-.252	-.151	.455*	.151	-.455*	13
68	.045	-.137	.411	-.045	.045	15
71	-.258	-.086	.258	.086	-.258	5
73	.081	.081	.244	.244	-.081	5
79	-.179	-.059	.299	-.179	-.419*	9
80	-.086	.258	.258	.086	-.430*	5
81	-.162	-.081	.244	.000	-.488*	12
83	-.111	.447*	.111	.000	-.223	10
86	.000	.000	.320	.160	-.160	6
87	.000	-.150	.150	.150	-.301	6
88	-.059	.179	.419	.059	.059	9
90	.070	.070	.212	.212	-.496*	7
91	.064	.452*	.323	.064	-.064	9
93	-.079	-.079	-.079	-.079	-.396*	5
94	.053	-.053	-.053	.373*	-.587*	11
96	-.126	.000	.126	.000	-.252	8
97	-.244	-.244	.244	.081	-.081	5
98	-.068	-.068	.068	.068	-.341	7
99	-.067	-.200	.200	.200	-.200	7
100	.000	.000	.112	.250	-.500*	10
101	.000	-.145	-.145	.291	-.437*	6

Note. Regression equations were calculated for teachers who had five or more grade discrepancies.

Teachers' grade discrepancies were coded -4, -3, -2, -1, 0, +1, +2, +3, or +4 according to whether a teacher's assigned grade was one, two, three, or four letter grades below or above the average of a student's academic performance.

Teacher's grade discrepancies were regressed on Student Informational Elements.

Table 5 (Cont'd.)

Influence of Non-achievement Factors on Secondary Teacher's Grading Practices: Standardized Regression Coefficients for Each Teacher (N=99)

Teacher	Student Informational Elements					Number of Discrepancies
	Gender	Effort	Behavior	Ability	Gr/Change	
108	.244	-.081	.081	.081	-.081	5
111	-.126	.378*	.504*	-.252	.000	8
112	-.079	.079	.237	-.237	-.396*	5
114	-.111	.000	.223	.111	-.447*	10
115	-.081	.081	.081	.081	-.406*	5
117	.000	.145	.000	-.145	-.291	6
118	.111	-.111	.111	.223	-.559*	10
119	-.042	-.042	-.042	.214	-.729	17
120	.079	.079	.237	-.237	-.237	5
121	-.068	-.068	.204	-.068	-.068	7
122	.079	-.237	.237	.396*	-.079	5
123	-.068	-.477*	.341*	.204	.204	7
124	-.064	-.064	.064	-.194	-.582*	9
125	.086	-.086	.258	.086	-.258	5
126	-.237	.118	.474*	.118	-.118	10
130	-.086	-.258	.086	-.086	-.430	5
132	-.067	-.067	.200	.200	-.067	7
133	-.049	.247	.347*	.148	.247	13
136	-.212	-.496*	-.070	-.212	-.212	7
138	-.050	-.252	.151	-.050	-.354	5
140	-.165	-.165	-.055	.055	-.497*	11
141	-.190	-.299	.027	-.027	-.190	18
142	-.059	.531*	.295	-.059	.059	9
148	.125	.000	.125	.000	-.500*	8
149	.134	.000	.404*	.269	-.134	8
150	.067	.468*	.334*	-.200	-.334*	7

Note. Regression equations were calculated for teachers who had five or more grade discrepancies.

Teachers' grade discrepancies were coded -4, -3, -2, -1, 0, +1, +2, +3, or +4 according to whether a teacher's assigned grade was one, two, three, or four letter grades below or above the average of a student's academic performance.

Teacher's grade discrepancies were regressed on Student Informational Elements.

Table 5 (Cont'd.)

Influence of Non-achievement Factors on Secondary Teacher's Grading Practices: Standardized Regression Coefficients for Each Teacher (N=99)

<u>Teacher</u>	<u>Student Informational Elements</u>					<u>Number of Discrepancies</u>
	Gender	Effort	Behavior	Ability	Gr/Change	
151	-.067	-.067	.067	.200	-.468*	7
153	.000	-.118	-.118	.356*	-.474*	12
154	-.258	.258	.430*	.258	-.086	5
155	-.070	-.070	.212	.354*	-.212	7
156	-.081	-.244	.081	.244	-.406*	5
158	-.244	-.244	-.081	.244	-.244	5
159	-.219	-.013	.395*	-.131	-.131	9
160	.000	.291	.437*	.000	.000	6
162	.150	.150	.301	.000	-.150	6
163	.144	.288	.000	-.144	-.288	6
164	-.081	.081	.081	.081	-.406*	5
165	.185	.185	.308	-.185	-.185	9
167	.000	.150	.301	-.150	.301	6
169	.000	-.112	.225	.112	.112	10
170	.000	.150	.150	.000	-.452*	6
172	-.129	.258	.387*	.129	.000	8
175	.081	-.244	.244	.244	.081	5
182	-.299*	.059	.538*	.179	-.179	9
185	-.060	-.060	.060	-.060	-.668*	11
187	.086	-.086	.258	.430*	.086	5
188	.301	.150	.150	.000	.000	6
189	-.258	-.086	.086	.086	-.430*	5

Note. Regression equations were calculated for teachers who had five or more grade discrepancies.

Teachers' grade discrepancies were coded -4, -3, -2, -1, 0, +1, +2, +3, or +4 according to whether a teacher's assigned grade was one, two, three, or four letter grades below or above the average of a student's academic performance.

Teacher's grade discrepancies were regressed on Student Informational Elements.

*p<.05

predictors of teachers' grade discrepancies have been identified with an asterisk. A summary of the number of teachers for whom a student characteristic factor was significant is provided in Table 6.

Most Frequently Occurring Significant Factors

While all five of the student characteristic factors were found to be a significant factor for at least one teacher, several factors were significant more frequently than others. The factor most frequently found significant in predicting grade discrepancies was "grade change." "Grade change" refers to whether grades reported in the vignettes for the average of four quizzes taken during the six-weeks were higher or lower than the six-week's test grade. In forty of the 99 teachers' regression equations (40%), grade change was a significant factor. Furthermore, in all of the regression equations where the grade change factor was significant the regression coefficients were negative, indicating teachers awarded higher grades when the quiz average was higher than the six-week's test grade and awarded lower grades when the six-week's test score was higher than the quiz average. Apparently teachers place more importance in the four quiz scores than they do on scores from a single six-week test.

Student behavior was the next most frequent factor found to be significant. It was found significant for 16 of the 99 teachers. In each case, the regression coefficient was positive, indicating higher grades associated

Table 6

Number of Teachers For Whom A Student Informational Factor Was a Significant Predictor* of Discrepancies Between Grades Awarded and Grades Based on Academic Performance

(N=99)

Gender	Effort	Behavior	Ability	Grade Change
1	10	16	7	40

Note. Regression equations were calculated for teachers who had five or more grade discrepancies.

Teachers' grade discrepancies were coded -4, -3, -2, -1, 0, +1, +2, +3, or +4 according to whether a teacher's assigned grade was one, two, three, or four letter grades below or above the average of a student's academic performance.

Teacher's grade discrepancies were regressed on Student Informational Elements.

* $p < .05$

with desirable behavior and lower grades associated with less desirable behavior.

Third in frequency of importance was student effort. Regression equations for 10 teachers had significant regression coefficients for this factor. Eight teachers had positive coefficients, while the coefficients for two teachers were negative. The positive coefficients indicate teachers "rewarding" high student effort with grade increases and "penalizing" low student effort with grade decreases. However, the two negative coefficients indicate just the opposite. Two teachers (one biology and one English) awarded students displaying low effort, increases in students' earned grades, and awarded students displaying high effort, grades below what had been earned on the tests. Closer inspection of the two teachers' grade discrepancies found neither teacher having grade discrepancies for the same vignette. While each teacher had seven grade discrepancies for the 32 vignettes, one awarded students grades lower than the grades implied in the vignettes in five vignettes where students were presented as demonstrating high effort. The other teacher awarded grades higher than those implied in six vignettes when students were depicted as demonstrating low effort. The only commonality found among the two teachers was both having an excess of twenty years of teaching experience.

Fourth in frequency of importance was student ability. Regression equations for 7 teachers had significant regression coefficients for the factor

student ability. All seven teachers' coefficients were positive, indicating teachers lowering the earned grade average of low ability students and raising the earned grade average of high ability students.

Finally, for one teacher the coefficient for the factor gender was significant. In this one case, the coefficient was negative, indicating lower grades awarded for female students while higher grades were awarded for male students. The teacher was female.

Policy Capturing Weights

In order that comparisons could be made concerning importance placed on the five informational cues by teachers in making grading decisions, policy capturing weights were calculated for each of the 99 teachers using the formula suggested by Hoffman (1960). Calculation of weights for each of the 99 teachers was as follows:

$$w_i = B_i r_i / R^2$$

w_i = weight for the i th attribute

B_i = standardized regression coefficient

r_i = correlation between the dependent variable and the i th attribute

Policy capturing weights were calculated in order for them to be compared both across teachers and across informational elements. The captured policies of teachers indicated substantial variation in importance assigned the informational elements. A comparison of teachers' policy

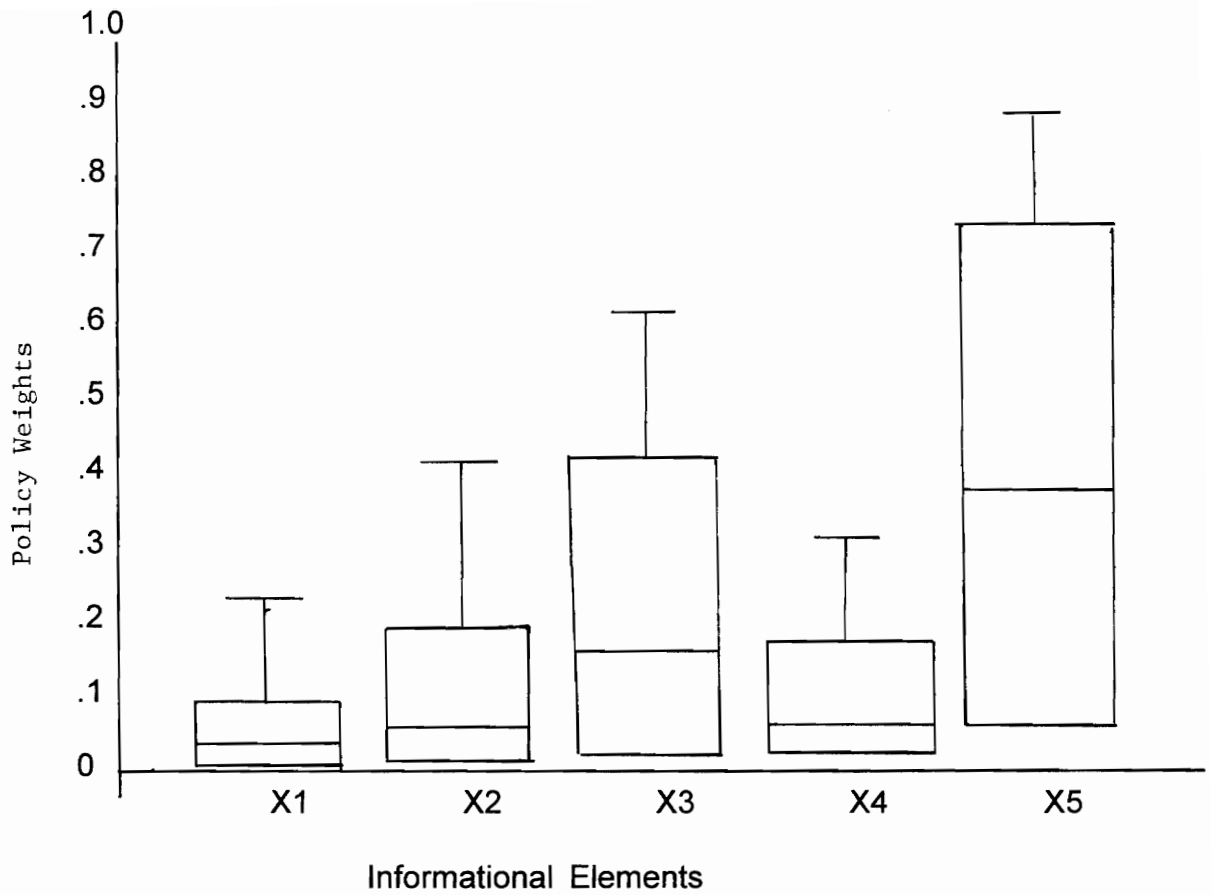
weighting of the informational elements is provided in boxplot charts in Figure 2.

The highest policy weights by teachers were assigned to grade change, receiving a median policy weight of 0.37 (range 0 to .96). Student behavior received the next highest median weight, 0.18 (range 0 to .87). Median weights for the remaining three factors were close to zero. Median weights were as follows: student ability, 0.07 (range 0 to .67); student effort, 0.06 (range 0 to .81); and student gender, 0.03 (range 0 to .69).

Variation in policy weighting for a particular factor has been interpreted as a lack of consistency in importance placed on the factor by raters (See e.g., Jaeger & Usher, 1991). However, in this study only two of the five informational elements seemed important to teachers in making their grading decisions. These were the factors grade change and student behavior. While the other three factors (effort, ability, and gender) showed much less variation in their policy capturing weights, this was more the result of their respective means being close to zero rather than consistency by teachers in considering them as important factors.

Teachers' Characteristics

In determining whether teacher characteristics could be used to predict the importance teachers placed on the five informational elements (i.e., student gender, effort, behavior, ability, and grade change), a secondary



X1 = Student Gender; X2 = Student Effort; X3 = Student Behavior;
 X4 = Student Ability; X5 = Grade Change during the grading period

Note. Bottom and top of boxes in figure occur at the 25th and 75th percentiles, respectively. The vertical lines extending above each box ends at the 90th percentile. Horizontal lines within each of the boxes represent position of the median.

Figure 2. Boxplot Charts of Policy Capturing Weights of Teachers Having Five or More Grade Discrepancies (N=99).

analysis of the data was conducted. It was anticipated that all 99 teachers having five or more grade discrepancies could be used as the sample. However, two of the teachers failed to provide information pertaining to their gender and years of teaching experience, therefore the sample was reduced to 97.

Five separate regression equations, each using a different dependent variable, were created. Dependent variables were teachers' standardized regression coefficients for the five informational elements (factors) student gender, effort, behavior, ability, and grade change. Regression coefficients were taken from each teacher's individual regression equation.

Each of the five regression equations created used the same independent variables which were four teacher characteristic variables. Each teacher had answered questions pertaining to their gender, educational level, number of years teaching, and subject area taught.

One of the variables, subject area taught, was recoded. Originally, the subject area variable consisted of seven categories however, history and government were combined into a single category, as were the subject areas foreign language and social studies. The resulting five teaching subjects were then entered into the regression equations using four dummy coded variables.

Stepwise regression analysis was used wherein the variable having the highest zero-order correlation with the dependent variable is entered first into

the regression equation. The order and presence of any other variable(s) entering the equation is based on those producing the greatest incremental change in the R-square, while taking into account any variable(s) already in the equation. It was decided to use (forward) stepwise regression since there was no reason to assume that the independent variables should be entered into equations in any particular order.

Two of the five regression equations created had significant multiple correlations. These were the equations using teachers' regression coefficients for the factors grade change and student behavior as dependent variables. In both equations, the same two teacher characteristic variables were found significant, teaching experience and the dummy coded variable representing teachers who taught either history or government subjects.

In the equation where coefficients for the grade change factor were used as the dependent variable, teaching experience was negatively related to importance placed by teachers on students' grade changes. The negative coefficient indicated that the more experience a teacher had the less importance they placed on students' grade changes. There was a positive relationship between the subject area taught variable, which had been dummy coded, and the importance placed on students' grade changes. This indicated that teachers who taught either history or government, rather than teachers in uncoded groups placed importance on students' grade changes.

In the other regression equation, using teachers' regression coefficients for student behavior as the dependent variable, teaching experience was again negatively related to importance placed on the student behavior, indicating the more experience a teacher had the less importance they placed on students' behavior. Again, teachers who taught history or government, as opposed to those who did not, were found to place importance on students' behavior. A comparison of changes in R-square's, standardized coefficients, and t-values of the factors for the two equations is provided in Table 7.

Frequencies of Grade Discrepancies by Vignettes

Grade discrepancies for each of the vignettes were totaled for all the participating teachers (N=192). A listing of the vignettes generating the most grade discrepancies and their corresponding factor level codings are provided in Table 8.

Two of the vignettes (#13 and #7) received grade discrepancies from over half the teachers in the study (105 and 98 discrepancies, respectively). In both of these vignettes the grade awarded was higher than the grade implied by the test scores. None of the teachers lowered any student's grade below the grade implied by the test scores. Similarities among the two vignettes consisted of students in both being male, exhibiting desirable behavior, and experiencing grade decreases (from their quiz average to their six-weeks test grade).

Table 7

Teachers' Characteristics That Predict* Importance Placed on the Non-achievement Factors: Student Behavior and Grade Change
(N=97)

Dependent Variable: Grade Change

<u>Variables</u>	<u>Standardized Estimate</u>	<u>T-Value*</u>	<u>R-Square Change</u>
Yrs-Tch	-.250	-2.7	.063
Sub4 (hist. & govt.)	.340	3.6	.113

R-Squared = .1636

Dependent Variable: Student Behavior

<u>Variables</u>	<u>Standardized Estimate</u>	<u>T-Value*</u>	<u>R-Square Change</u>
Yrs-Tch	-.280	-2.9	.076
Sub4 (hist. & govt.)	.260	2.7	.069

R-Squared = .1357

* $p < .05$

Table 8

Number of Discrepancies, Percentage, Corresponding Vignette Number, and Profile Coding* of the Most Positive and Negative Grade Discrepancies by Teachers

(N=192)

Positive Discrepancies

<u>Discrepancies</u>	<u>Percentage</u>	<u>Vignette</u>	<u>Profile Coding</u>
105**	54.6%	#13	(1,2,2,1,1)
98**	51%	#7	(1,1,2,2,1)
92	47.9%	#19	(2,1,1,2,1)
57	29.6%	#29	(2,2,2,1,1)
51	26.5%	#15	(1,2,2,2,1)
43**	22.3%	#16	(1,2,2,2,2)

Negative Discrepancies

<u>Discrepancies</u>	<u>Percentage</u>	<u>Vignette</u>	<u>Profile Coding</u>
56***	29%	#8	(1,1,2,2,2)
50***	26%	#28	(2,2,1,2,2)
32***	16.6%	#26	(2,2,1,1,2)
32	16.6%	#10	(1,2,1,1,2)
26	13.5%	#6	(1,1,2,1,2)

* Order and Value of Profile Coding:
 first, gender (1 = male, 2 = female);
 second, effort (1 = low, 2 = high);
 third, behavior (1 = unacceptable, 2 = desirable);
 fourth, ability (1 = low, 2 = high); and
 fifth, grade change (1 = decrease, 2 = increase; from quiz average grade to six-week test grade).

** All grade discrepancies for the vignette were positive.

*** All grade discrepancies for the vignette were negative.

Two other vignettes (#8 and #28) were distinct in that they received the highest numbers of negative grade discrepancies from teachers (56 and 50 discrepancies, respectively). In both vignettes, the grades awarded by teachers, when discrepancies occurred, were always below the grades implied by the test scores in the vignettes. Similarities among the vignettes included students in both being depicted as having high ability and a quiz average which was lower than their six-weeks test grade.

Further analysis was conducted regarding the textual presentation and organization of the informational elements in the four vignettes (#13, #7, #8, and #28). Vignette number 13, which had received the highest number of positive grade discrepancies (105), contained five statements promoting the "goodness" of the male student depicted to the extent that the student could have been categorized as a "teacher's pet". Three statements presented desirable classroom behavior and two presented high effort. The student's grades were presented as decreasing from a B average for the four quizzes to a C- on the six-week's test.

Vignette number 7, which had received the next highest number of positive discrepancies (98), presented a male football star. There were two statements referring to the student's low effort regarding his classwork. The remaining depiction presented several statements referring to the student's desirable behavior and high ability. The student's grades were presented as

decreasing from a C average for the four quizzes to a D- on the six-week's test.

Vignette number 8, which had received the next to highest number of negative grade discrepancies (56), presented in its first sentence a male student never doing any more than he absolutely had to in class (low effort). However, in the second sentence the student is depicted as having acceptable behavior, hurrying each day from his gym class to get to your class on time. Another statement was presented regarding the student's low effort, "seldom has his homework done...He offers excuses." The vignette ends with two positive informational elements (high ability and grade increases). The student's grades were presented as increasing from a D+ average on the four quizzes to a C on the six-week's test.

The other vignette, number 28, which had received the highest number of negative grade discrepancies (50), presented five negative phrases (e.g., absent from school a lot; makes a lot of noise) or words (e.g., tardy) regarding unacceptable behavior for the female student depicted. All five negative statements were within the first four sentences of the vignette. Also, included among the statements was a depiction of the student as cocky with a "so what" attitude concerning her behavior. So much unacceptable behavior at the beginning of the vignette perhaps tainted the remaining informational elements presented which were all positive (high effort, high ability, and grade increases). The student's grades were presented as

increasing from a C- average on the four quizzes to an A- on the six-week's test. Content and organization of the four vignettes is provided in Figure 3.

Teachers' Comments

Teachers were invited to comment regarding the vignettes or problems they may have encountered in doing the exercise. Therefore, any comments teachers made were entirely voluntary. Sixty-nine (36%) of the 192 teachers provided at least one comment, many providing several.

Comments of the teachers are presented in Appendix D. Included along with each teacher's comment(s) is the teacher's subject area, grade level, years of teaching experience, number of grade discrepancies, and, in the case of there being five or more grade discrepancies, the significant factors from the teacher's individual regression equation. This information was included to provide a profile of teachers offering each comment so as to judge whether the teachers' grading practices are consistent with their comments. For example, one teacher with 29 years of teaching experience, commented that the exercise was very interesting and that she had "changed (her) my mind (and her responses) after re-reading the instructions and the vignettes" such that, her original grade selections for the vignettes had contained over 15 grade discrepancies, but after considering her grading philosophy she changed her grades assigned to the vignettes, which then resulted in only five discrepancies.

Positive Discrepancies

Vignette #13; **105 Discrepancies (54.6%)**

Profile levels (1,2,2,1,1)

- male
- high effort
- desirable behavior
- low ability
- decrease in grades (B to C-)

Donnie does a lot of little extra things for you, like running errands to the office and helping straighten up the desks after class.

You and Donnie get along very well.

Based on his classwork and the opinion of other teachers you judge he would be placed in the low ability range.

However, even though his ability is on the low side, he works hard in class and seems genuinely interested.

He even asks for extra work to do outside of class.

On the four quizzes you gave during the six-weeks he had a B average, but got a C- on your six-week test.

What grade you would assign Donnie for the six-weeks?

A, B, C, D, F

Figure 3. Content and Organization of Vignettes Receiving Highest Number of Positive and Negative Grade Discrepancies from Teachers (N = 192)

Positive Discrepancies

Vignette #7; **98** **Discrepancies (51%)**

Profile levels (1,1,2,2,1)

- male
- low effort
- desirable behavior
- high ability
- decrease in grades (C to D-)

Ronnie is the kind of student that only does what he "has to" academically, but is enthusiastic about sports and one of the stars on the football team.

You have talked to him about applying himself in class because you think he is smart enough to make it in college.

You have told him that he could make better grades if he applied himself, but he would rather pursue his interests in sports and stock car racing.

He is seldom prepared for class, but has never given you any trouble in terms of his class behavior.

He had a C average on the four quizzes you gave, then got a D- on the six-week test.

What grade you would assign Ronnie for the six-weeks?
A, B, C, D, F

Figure 3 (Cont'd.)

Negative Discrepancies

Vignette #8; **56 Discrepancies (29%)**

Profile levels (1,1,2,2,2)

- male
- low effort
- desirable behavior
- high ability
- increase in grades (D+ to C)

Robert never wants to do any more than he absolutely has to in your class.

He has gym just before your class and often has to hurry to get there on time, but he always makes it.

However, he seldom has his homework done, always wanting more time to complete it.

He offers excuses such as he had to go visit a sick relative or left his book at school, etc.

He had a D+ average on the four quizzes you gave during the six-weeks and a C on the six-week test.

You think he's much smarter than his grades indicate.

Other teachers have told you that he does very well in their classes.

What grade you would assign Robert for the six-weeks?

A, B, C, D, F

Figure 3 (Cont'd.)

Negative Discrepancies

Vignette #28; **50 Discrepancies (26%)**

Profile levels (2,2,1,2,2)

- female
- high effort
- unacceptable behavior
- high ability
- increase in grades (C- to A-)

Thelma is absent from school a lot and often tardy to your class when in attendance.

She missed eight days the last six-weeks.

It's bad enough that she is tardy but she also makes a lot of noise when she comes in.

When you have spoken to her about her tardiness and her disrupting class, her attitude has always been "so what".

You've seen her permanent record and know that she placed in the top quartile on a standardized ability measure.

She is a talker and likes to take part in the discussions you have in class.

During the last six-weeks she had a C- average on the four quizzes you gave (she was in school for all of them) and had an A- on her six-weeks test.

What grade you would assign Thelma for the six-weeks?
A, B, C, D, F

Figure 3 (Cont'd.)

Teachers' comments were first coded into 18 different categories, consisting, for example, of such things as not enough information provided within the vignettes, problems in calculating grade averages, and including student's daily grades as part of grading criteria. However, it was found that four of the 18 categories received the most frequent comment from the teachers. Ranking first in frequency was the comment that "student's grades should consist solely of academic achievement." Twenty-six of the teachers (13.5%) stated that they based grading on academic achievement only. To determine whether or not teachers did what they said they did, these 26 teachers were more closely examined with regard to their number of grade discrepancies. Of the 26 teachers, eight teachers (30.7%) had 5 or more grade discrepancies. Four (50%) of these teachers had an excess of 10 discrepancies and one had 16 discrepancies.

The second most frequent comment concerned the fact that homework completion and class participation were not mentioned in the information presented in the vignettes. Thirteen of the teachers (7%) commented in some way on the importance of using homework completion and class participation in awarding grades.

Two categories were commented on by seven teachers (4%). The categories were "that student behavior should not be a factor in student grading" and "that close calls in grading should be influenced by a student's effort, standardized test scoring, or other circumstances."

Ranking fourth with five comments was "that student's grades should be based on such things as the student's work, participation in class, and improvement." Also, included in the category were "things the student does best" and "the student's effort."

Perhaps the most revealing finding from teachers' comments was that so few teachers (13%) offered a comment concerning the importance of basing grades on solely students' academic performance. Disturbingly, when those teachers who did offer such comment were given the opportunity, more than one-fourth of them awarded five or more grades (out of the 32 possibilities) which were different than the grades implied in the vignettes.

Conclusions

Secondary school teachers readily acknowledge including extraneous non-achievement related factors in grades they award students (Brookhart, 1991; Frary et al., 1993; Nova & Loyd, 1992). Defending their actions, teachers point out that it is often necessary for them to consider the impact a grade may have on a student, such as in creating undesirable social consequences (Brookhart, 1991).

However, it should be pointed out that if students' grades are based on factors other than a student's academic performance, there is room for ambiguity and misinterpretation. Grading practices attempting to moderate the impact grades may have on a student by including non-achievement factors raises the question of grade interpretation. Measurement specialists agree that grades should be based solely on a student's academic performance (Hopkins et al., 1990). However, evidence is that high school teachers fail to heed to such advice, knowingly including non-achievement factors in their grading (Brookhart, 1991; Frary et al., 1993; Nova & Loyd, 1992).

The present study attempted to identify the relative influence non-achievement factors had on secondary teachers grading practices. It specifically attempted to predict on the basis of non-achievement related factors the discrepancy between grades awarded by teachers to hypothetical

student profiles and grades implied in the profiles if based solely on academic achievement.

In order to analyze the existence of any systematic use by teachers of non-achievement related factors (i.e., gender, effort, behavior, ability, and grade change), it was necessary for sufficient variation to be present between teachers' awarded grades and grades implied in the vignettes. Therefore, the number of grade discrepancies for teachers were restricted to at least five or more. While this restriction provided sufficient variation on the criterion variable, it reduced the number of teachers in the sample by almost half (51.5%), from 192 teachers down to 99. It should be pointed out that of the 192 teachers with fewer than five grade discrepancies, thirteen (6.8%) of them failed to have any grade discrepancies for the 32 vignettes, indicating that some teachers base their awarded grades solely on students' academic performance.

While it could be considered that a reason for lack of variation on the criterion variable was that teachers in the study modified their natural way of grading to suit the study's purpose, there was no evidence supporting such claim. No indication was found from teachers' written or verbal comments after completing the exercise that they understood how the vignettes in the exercise had been constructed or were being scored.

Lack of variability has not been presented as a problem in other studies dealing with non-achievement factors in teachers' grading (Manke &

Loyd, 1990; Nova & Loyd, 1992), nor have studies using policy capturing methodology mentioned lack of variation on the criterion variable as a limiting factor (For examples of policy capturing studies, see the review by Hobson & Gibson, 1983). It was, therefore, surprising that many of the teachers in the sample had few grade discrepancies because during pilot testing sufficient numbers of grade discrepancies had been generated by teachers.

Because of a concern regarding differences existing in the number of grade discrepancies among the teachers, teachers were separated into groups of zero, one-to-four, and five-or-greater grade discrepancies and compared. The three groups were not found to differ significantly on any of the teacher characteristic variables, which included amount of teaching experience, subject area, and gender. Additionally, neither groups' discrepancies seemed to center on any particular vignettes. It was concluded that teachers' grade discrepancies could be attributed to the personal grading practices of the teachers themselves. This finding is supported by Borko and Caldwell's (1982) study which found teachers' decision-making independent and individually based.

In gathering the sample of teachers to participate in the grading exercise, it was assumed that high school teachers teaching academic subjects either considered non-achievement factors in awarding students' grades or did not. It was felt teachers' responses to the exercise would not be influenced by such things as their demographic region or size of school.

Therefore, teachers in the sample were not randomly selected. The sample of teachers obtained should be considered representative for the purpose of this study which was to identify the relative influence non-achievement factors had on high school teachers' grading practices.

Findings of the present study suggest that one out of every two high school teachers, who teach academic subjects, consider two non-achievement related factors, grade change and student behavior, in their grading. The grade change factor refers to whether the grade average of the four quizzes taken during the six-weeks was higher or lower than the six-week's test grade. Student behavior refers to whether a student's classroom behavior was unacceptable or desirable. The following conclusions are drawn from the 99 teachers in the sample who had five or more grade discrepancies for the 32 vignettes.

Effect of Grade Change

Grade change was the factor receiving the highest consideration among teachers. More specifically, this finding suggests that teachers give students' quiz average grades more weight than students' six-week test grades. This could, perhaps, be because teachers feel that an average of four quiz grades is a better indicator of student performance than a single test grade. Certainly, several quiz grades averaged together provides a student an opportunity to "pull-up" one or more low grades.

This explanation is conceivable. However, it remains incongruous for teachers to award students grades corresponding to their quiz average, disregarding a student's performance on a six-week test.

Effect of Student Behavior

Student behavior was the other factor teachers considered most when assigning grades. Teachers are influenced by student behavior in their grading, awarding higher grades to students exhibiting desirable behavior and lower grades to students exhibiting less desirable behavior.

The kinds of behavior that were viewed favorably by teachers and affected teachers' grading positively were students: (1) behaving in such a manner that was conducive to learning in class, such as respecting other student's rights and being polite; (2) being considerate of the teacher and the other students in the class; (3) doing extra things during class, like running errands, straightening desks, or washing the board; (4) making extra effort to be on time to class; and/or (5) being trustworthy. Conversely, behaviors which were viewed negatively by teachers in assigning grades were students: (1) talking-out without permission; (2) horsing around; (3) being disrespectful; (4) using profanity; (5) cheating or stealing; (6) tardiness to class; and/or (7) breaking school rules and receiving a detention notice from a teacher.

Previous studies have found teachers saying that in some instances a student's behavior was considered when grades were awarded (Agnew,

1985; Frary et al., 1993; Nova & Loyd, 1992). However, the importance of student behavior, at least in the studies reviewed, was always preceded by student effort. Findings from this study suggest teachers place more importance in their grading on student behavior than on student effort.

An additional finding from this study is that more experienced teachers place less importance on grade change or student behavior in their grading than less experienced teachers. This finding suggests that teachers tend to become more academic performance oriented regarding their grading from the on-the-job experience of teaching, pointing to a possible inadequacy in teacher education programs regarding teacher preparation in grading policies and practices.

In the grading practice studies reviewed, teacher experience was not addressed in terms of its relationship with influence of non-achievement factors in teachers' grading practices. From findings in this study, further investigation is suggested regarding the relationship.

When grade discrepancies from all 192 teachers in the study were totalled for each of the 32 vignettes, four vignettes were found receiving the greatest number of grade discrepancies from teachers. For two of these vignettes, discrepancies were observed for over half of the participants. In each of the vignettes, teachers' grade discrepancies were always positive, indicating that their awarded grade was, in every instance, higher than the grade implied by the test scores in the vignettes. Both vignettes depicted a

male student exhibiting desirable classroom behavior; one described a student as a "teacher's pet" and the other referred to the student as a "football star." In both vignettes, the student's quiz average was higher than their six-week test grade. These results suggest teachers rewarding preferential classroom behavior and tending to make special concessions for athletes.

Approximately one-fourth of the teachers had a negative grade discrepancy for two vignettes. In both vignettes there were only negative grade discrepancies from teachers. Similarities and identification of students in these profiles were not as easily discerned as the two vignettes having students receiving higher grades. Students were presented in both as having high ability and a lower quiz average grade than six-week test grade. One vignette presented several negative statements at its beginning with regard to unacceptable behavior by the student. In the other, positive statements regarding the student's behavior and ability were presented. The negative discrepancies for these two vignettes may be explained by teachers focusing on unacceptable behavior and poor effort, awarding a lower grade than implied in the vignette.

Implications and Recommendations for Future Research

The following insights and suggestions are offered.

1. Findings from this study indicate that one out of every two secondary school teachers teaching academic subjects award five students in each of their classes grades which include non-achievement related factors.

Furthermore, teachers who include non-achievement related factors in their awarded grades often teach as many as six classes per day, which means thirty students receiving grades either higher or lower than actually earned.

In an even broader perspective, take for example a typical secondary school having ten teachers teaching academic subjects. If five of these teachers fall into the category of awarding grades based on factors other than solely students' academic performance, there would be a total of one hundred and fifty students in the school receiving grades based on factors other than their academic performance. This example illustrates the extent of the influence of non-achievement related factors on grades in secondary schools.

2. Vignettes depicting academically at-risk students were found to receive grades from approximately one-fourth of the teachers in the sample of 192, which were at least one letter grade lower than grades which they had earned. Such occurrence is especially noteworthy since at-risk students report low grades to be one of the major reasons they dropout of school.

In the study, three vignettes depicted students who could be characterized as potentially at-risk. In two of these three vignettes, over one-

fourth of the 192 teachers awarded grades to the students which were lower than the grades the students had actually earned. One vignette presented a student who had earned a grade of C. The student's grade was lowered to a grade of D by 55 teachers and to an F by one teacher. Another vignette presented a student who had earned a grade of B. This student received a grade of C from 43 teachers and a grade of F from three teachers. Finally, in another vignette almost one-fifth of the teachers (32) awarded the student a failing grade of F when the student had earned a passing grade of D.

Teachers lowering students' grades presented in these three scenarios can be damaging to any student, much less the at-risk student. However, where the lowered grade falls on the grading scale in each of the scenarios is of particular interest. In one scenario, a potentially honor role grade of B was lowered to a grade of C. In another scenario, a grade of C, which is satisfactory for extra curricular participation in many schools, was lowered to a grade of D. Finally, one scenario depicted a struggling student earning a passing grade of D. Lowering the grade, in this case, could result in a demoralizing F for the student.

Certainly grades can impact any student's self-esteem or self-worth, but they have special significance for the at-risk student who perhaps only needs a nudge to dropout and never be seen again except, perhaps, on an unemployment or welfare roster. If the grades awarded to these students represent something other than the students' academic performance, any

belief they may have in the "system's" fairness or equality with concern to participating students may be jeopardized.

3. The gestures, actions (note passing), and verbal comments made by teachers while (and after) participating in the grading exercise, and also their written comments, strongly suggest that the exercise prompted them to think about their personal grading practices. There is no way in knowing what impact the exercise had on teachers in this regard. On the down side, the discussion among teachers during the exercise, while a limitation of the methodology, may explain the failure of some to have many grade discrepancies. For example, one teacher commented that she had "gone back" and changed several students' grades after completing the exercise and rethinking her own grading policy.

4. The majority of comments made by teachers expressed a belief that student's grades should consist solely of academic achievement. However, in this study, many of the teachers making such comments were found to have moderate to large numbers of grade discrepancies. This raises a question of accuracy regarding surveys eliciting information from teachers concerning the non-achievement factors they consider in awarding grades. Further investigation is suggested with regard to gaining insight into what teachers actually do with regard to grading, rather than what they say they do.

5. Even though grade discrepancy differences among teachers were examined with regard to a few teacher characteristics, it is possible that other

teacher variables such as teacher training, teaching styles, or classroom settings could relate more strongly to teacher's grade discrepancy differences. It is suggested that such variables be included in any replication of this study.

6. It was not the intention of the present study to include any additional student characteristic factors in the vignettes other than the five selected.

However, from this study's findings two additional factors seem to also influence teachers' grading, preferential classroom behavior (teachers pet) and student status (football star). Further investigation is suggested regarding these factors along with textual presentation of factors in simulations, which also could have been a contributing factor.

7. Teachers placed less importance on non-achievement related factors in making grading decisions the more teaching experience they had. This finding suggests an evaluation of the content of teacher preparation and inservice grading instruction programs regarding grading practices. Also, more experienced teachers placing less importance on non-achievement factors in grading supports master teacher and mentoring type programs where experienced teachers have the opportunity to demonstrate and advise less experienced teachers concerning grading practices.

8. Finally, teachers were found influenced, mainly, by two factors in awarding students grades, grade change and student behavior. It would seem prudent to direct attention, in teacher preparation and inservice programs, toward off-setting what seems to be an inclination of teachers to

consider these two factors in awarding students' grades. Also, further investigation is suggested regarding why teachers are inclined to include these non-achievement related factors more than others.

References

- Agnew, E. (1985). The grading policies and practices of high school teachers. ED259022. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Airasian, P. W. (1991). Classroom assessment. New York: McGraw-Hill.
- Anderson, B. L. (1977). Differences in teachers' judgement policies for varying numbers of verbal and numerical cues. Behavior and Human Performance, 19, 68-88.
- Barnes, S. (1985). A study of classroom pupil evaluation: the missing link in teacher education. Journal of Teacher Education, 36(4), 46-49.
- Borko, H., & Cadwell, J. (1982). Individual differences in teachers' decision strategies: An investigation of classroom organization and management decisions. Journal of Educational Psychology, 74, 598-610.
- Brookhart, S. M. (1991). Grading practices and validity. Educational Measurement: Issues and Practices, 10(1), 35-36.
- Brookhart, S. M. (1994). Teachers' grading: Practices and theory. Applied Measurement in Education, 7(4), 279-301.
- Feldman, D. C., & Arnold, H. J. (1978). Position choice: Comparing the importance of organizational and job factors. Journal of Applied Psychology, 63, 706-710.
- Frary, R. B., Cross, L. H., & Weber, L. J. (1993). Testing and grading practices and opinions of secondary teachers of academic subjects:

- Implications for instruction in measurement. Educational Measurement: Issues and Practice, 12, 23-30.
- Hobson, C. J., & Gibson, F. W. (1983). Policy capturing as an approach to understanding and improving performance appraisal: A review of the literature. Academy of Management Review, 8(4), 640-649.
- Hobson, C. J., Mendel, R. M., & Gibson, F. W. (1981). Clarifying performance appraisal criteria. Organizational Behavior and Human Performance, 28, 164-188.
- Hoffman, P. J. (1960). The paramorphic representation of clinical judgement. Psychological Bulletin, 57, 116-131.
- Hopkins, K. D., Stanley, J. C., & Hopkins, B. R. (1990). Educational and Psychological Measurement and Evaluation. Englewood Cliffs, NJ: Prentice-Hall.
- Jaeger, R. M., & Usher, C. H. (1991). Alternative procedures for integrating multidimensional evaluations of schools: An experimental comparison, Presented at the 199? Annual Meeting of the American Educational Research Association, Chicago, IL.
- Laughlin, J. E. (1978). Comment on "estimating coefficients in linear models: it don't make no nevermind. Psychological Bulletin, 85(2), 247-253.
- Manke, M. P., & Loyd, B. H. (1990, April). An investigation of non-achievement-related factors influencing teachers' grading practices.

Paper presented at the Annual Meeting of the National Council on Measurement in Education, Boston.

Nava, J. F., & Loyd, B. H. (1992, April). An investigation of achievement and nonachievement criteria in elementary and secondary school grading.

ED346145. Paper presented at the American Educational Research association, San Francisco.

Oskamp, S. (1967). Clinical judgement from the MMPI: Simple or complex? Journal of Clinical Psychology, 23, 411-415.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1977). Behavioral decision theory. Annual Review of Psychology, 28, 1-39.

Slovic, P., & Lichtenstein, S. (1977). Comparison of Bayesian and regression approaches to the study of information processing in judgement. Organizational Behavior and Human Performance, 6, 649-744.

Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgement. Organizational Behavior and Human Performance, 6, 649-744.

Stahl, M. J., & Zimmerer, T. W. (1977). Modeling strategic acquisition policies: A simulation of executives' acquisition decisions. Academy of Management Journal, 27, 369-383.

- Stiggins, R. J., Frisbie, D. A., & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. Educational Measurement: Issues and Practices, 8(2), 5-14.
- Stumpf, S. A., & London, M. (1981). Capturing rater policies in evaluating candidates for promotion. Academy of Management Journal, 24, 752-766.
- Taylor, R. L., & Wilsted, W. D. (1974). Capturing judgement policies: A field study of performance appraisal. Academy of Management Journal, 17, 440-449.
- Terwilliger, J. S. (1989). Classroom standard setting and grading practices. Educational Measurement: Issues and Practice, 8(2), 15-19.
- Tomchin, E. M. (1989). Teachers beliefs about grade retention. Dissertation Abstracts.
- Zedeck, S., & Kafry, D. (1977). Capturing rater policies for processing evaluation data. Organizational Behavior and Human Performance, 18, 269-294.

APPENDIX A

Time Line of Data Collection

(All school districts are in Virginia)

5-16-94 through 5-20-94, Pilot Study - Montgomery County Schools
(12 participants)

1st Wave Data Collection

5-17-94, Six letters sent to district superintendents requesting permission to invite teachers in their districts to participate in the study. Also, an additional division superintendent contacted (6-3-94).

6-10-94, Permission received from all seven school districts.

6-13-94 to 6-28-94, Data collected from seven school districts.

<u>School Division</u>	<u># Teachers Contacted</u>	<u># Exercises Returned</u> (Academic Teachers)
Wise County	30 mailed	19
Colonial Heights	20 mailed	7
Giles County		
- Giles H.S.	3 individ. contacted	3
- Narrows H.S.	2 individ. contacted	2
Montgomery County		
- AHS	10 tch. mailboxes	2
- BHS	15 tch. mailboxes	4
- SHS	10 tch. mailboxes	4
Pulaski County	30 tch. mailboxes	6
Fluvanna County	10 mailed	2
Rockbridge County	20 mailed	3
	<hr/>	<hr/>
	Total 150	52

Time Line of Data Collection (cont.)

2nd Wave Data Collection

6-20-94, Fifteen letters sent to district superintendents requesting permission to invite teachers to participate in study.

8-4-94, Permission received from 5 school districts.

8-18-94 to 8-28-94, Data collected from 5 school districts.

<u>School Division</u>	<u># Teachers Contacted</u>	<u># Exercises Returned</u> (Academic Teachers)
Smyth County		
- Northwood	35 teachers' meeting	13
- Marion	65 teachers' meeting	36
Grayson County	50 teachers' meeting	31
Radford City	50 tch. mailboxes	3
Botetourt County		
- James River	45 teachers' meeting	21
Carroll County	50 teachers' meeting	31
	_____	_____
Total	457	140

1st and 2nd Wave Data Collection Combined

<u>School Division</u>	<u># Teachers Contacted</u>	<u># Exercises Returned</u> (Academic Teachers)
12 Districts	607	192

APPENDIX B

Factor Levels and Coding

Thirty-two vignettes, or student profiles, using combinations of five informational elements were constructed. The informational elements (factors) used in the study and their respective codings include:

- (1) Sex (1=male, 2=female),
- (2) Effort (1=low, 2=high),
- (3) Behavior (1=unacceptable, 2=desirable),
- (4) Ability (1=low, 2=high), and
- (5) Test score improvement (grade change) - change from quiz average grade to test grade (1=decrease, 2=increase).

Presented at the top of each of the vignettes shown in Figure B-1, in parentheses, are the specific levels of the factors used in the simulation. Factor level presentations for the vignettes are ordered and coded as shown above.

Interpretations of Factor Levels

Below are the interpretations of the factors used in creating student profiles:

Sex. Male or female.

Effort. A measure of how much effort a student puts forth regarding class.

Low effort is interpreted as the student:

- (1) not doing at least 50% of the homework assigned for the class;
- (2) not being prepared for class by doing such things as assigned reading, projects, etc.;
- (3) not participating in class by doing such things as asking questions, showing interest, etc.;
- (4) not requesting extra credit work to do; and
- (5) only doing what they "have too" academically.

High effort is interpreted as the student:

- (1) doing most or all the homework for the class;
- (2) being prepared for class by doing such things as assigned reading, projects, etc.;
- (3) participating in class discussions, asking questions, showing interest, etc.; and
- (4) requesting extra credit work to do.

Behavior. A student's conduct in class.

Unacceptable behavior is interpreted as the student misbehaving in class by:

- (1) talking-out without permission;
- (2) horsing around;
- (3) being disrespectful;
- (4) using profanity;
- (5) cheating or stealing;
- (6) tardiness to class; and/or
- (7) breaking school rules and receiving a detention notice from a teacher.

Desirable behavior is interpreted as the student:

- (1) behaving in such a manner that is conducive to learning in class, such as respecting other student's rights and being polite;
- (2) being considerate of the teacher and the other students in the class;
- (3) doing extra things during class, like running errands, straightening desks, or washing the board;
- (4) making extra effort to be on time to class; and/or
- (5) being trustworthy.

Ability.

A measure of a student's overall ability, is presented as either:

- (1) a student's last standardized ability test score, in which a percentile score is given (scores below the 50th %ile are considered **low** ability and scores above the 50th %ile **high** ability), or
- (2) the teacher's knowledge of the student's classwork, classroom tests, and/or observation such that in their opinion the student exhibits either **low** or **high** ability.

The thirty-two vignettes are provided in Figure B-1.

Test Score Improvement (Grade Change).

Each of the 32 vignettes depicts a student acquiring two academic achievement grades, a quiz average grade (average of four quizzes taken during the six-weeks) and a six-week test grade. These two grades were varied for each student profile to show either an increase or decrease from the student's quiz average to six-week test grade.

The following grading scale and numeric values were used in creating grading simulations: A = 4, A- = 3.7, B+ = 3.3, B = 3, B- = 2.7, C+ = 2.3, C = 2, C- = 1.7, D+ = 1.3, D = 1, D- = .7, and F = 0. Grade increases or decreases from students' quiz average to six-week's test in the vignettes were based on specific factor level combinations in a vignette and on varying the amount increase or decrease in students' grades. In varying the amount of change in grades it was felt the change should be distributed among the following "letter grade differences":

- (1) one-half grade difference (0.3, e.g. D to D+),
- (2) one grade difference (0.7, e.g. B to A-),
- (3) one and one-half grade difference (1.3, e.g. C to B+; and 1.7, e.g., C to A-),
- (4) two grade difference (2.0, e.g., C- to A-; and 2.3, e.g. C- to A), and
- (5) two and one-half grade difference (2.7, e.g., F to B-).

Table B-1 provides the distribution of the grade differences.

Different Vignette Presentations

It was decided to create three different forms of the exercise, or order of vignette presentation, to guard against teachers making comparisons of

each others' responses. The order of vignette presentations for the three forms is provided in Table B-2.

Teachers' Directions for Completed the Grading Exercise

Attached to each packet of vignettes were a set of directions.

Directions asked teachers to read each of the vignettes and assign a grade to the hypothetical student depicted. Teachers were also asked to respond to the vignettes "in terms of one of the academic classes you are actually teaching this term" and record that information along with the grade level.

Along with directions, teachers were asked five questions pertaining to how many years the teacher had been teaching, highest level of educational attainment, gender, ethnicity, and subjects and grade levels they were currently teaching. Directions and questions are provided in Figure B-2.

#1 (1,1,1,1,1)

Because of his talking, you have had to change Sam's seating in your class several times. Twice, during the past six-weeks, he disturbed class so much by talking out and not being quiet when you asked that you had no choice but to send him out of class and later write him up. You referred him, earlier in the year, to be tested for perhaps a learning disability, but the school psychologist said he had been tested last year and was not learning disabled. His records show that on the last standardized ability test he took he was at the 40th percentile. Sam seldom does any homework, but managed to have a C- average on the four quizzes you gave during the six-weeks. However, he failed your six-week test. What grade you would assign Sam for the six-weeks? A, B, C, D, F

#2 (1,1,1,1,2)

John, a student in your class, had a D average on four quizzes you gave during the six-weeks. However, his grades improved somewhat during the grading period because he had a D+ on your final test. He is a little wiry student with shoulder length black hair which is never combed. Of all the students you have had, he has the most irritating laugh of any of them. In class, he is constant energy, fidgeting around, laughing uncontrollably at times, and a behavior problem day-after-day. His other teachers and you agree from his classwork, quizzes, and tests that he has below average ability. He seldom brings paper and pencil to class. What grade you would assign John for the six-weeks? A, B, C, D, F

#3 (1,1,1,2,1)

You can hear his voice just outside your room and know one of your more "difficult to manage" students, Chris, is at school. Chris usually dresses mostly in black, wears dirty jeans and motorcycle boots, and has a chain on his wallet. Recently, you had to write him up when he disrupted class by showing off his new tatoo. Chris seldom, if ever, participates in class. However, he seems to be rather smart. He has a knack for remembering things you say in class. If you stay after him, he can do good work. It surprises you sometimes how well he does. For example, during the past six-weeks he had a B average on the four quizzes you gave, but then got a D- on your six-week test. What grade would assign Chris for the six-weeks? A, B, C, D, F

Figure B-1. Vignettes and Factor Level Combinations

#4 (1,1,1,2,2)

Joe works on his father's dairy farm everyday after school. He is in your sixth period class. He has told you that working on the farm is what he knows he wants to do and doesn't see why he should finish school. He seldom does any homework for your class. You have talked to his parents and know they don't want him to dropout of school. Joe sometimes falls asleep in class. However, you know he's got the ability to do more than he does because his permanent record shows that he scored at the 88th percentile on the last administered standardized ability test. He had a C on the six-week test you gave and got a C+ on the four quizzes given during the grading period. What grade you would assign Joe for the six-weeks? A, B, C, D, F

#5 (1,1,2,1,1)

You have tried to get William more involved in class by calling on him with easy questions, but he still only participates when you force him to. He used to help you with some of the slower students in class. However, he recently stopped doing his homework. You happen to know that his parents are going through a divorce and that he has been living with his grandparents. His grades also dropped during the last six-weeks. He had a quiz average of C- (on four quizzes), but got an F on your six-week test. When you asked him what happened on the test he just shrugged his shoulders. On the most recent standardized ability test he took, he scored at the 45th percentile. What grade would assign William for the six-weeks? A, B, C, D, F

#6 (1,1,2,1,2)

Larry works at a local grocery after school. He told you that he has to work to help support his family. Larry speaks and understands English very well even though you know another language is spoken in his home. He is a nice boy, cooperative and quiet, helps you by washing the board sometimes, but seldom participates in class except when you call on him. Another teacher told you they had seen is permanent record and that he had scored low on a standardized ability test he took. While you think Larry's ability is probably on the low side to, you wonder about the cultural bias' of standardized tests. During the last six-weeks you gave four quizzes. Larry had a C- average on the quizzes and got an A- on his six-week test. What grade you would assign Juan for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd). Vignettes and Factor Level Combinations

#7 (1,1,2,2,1)

Ronnie is the kind of student that only does what he "has to" academically, but is enthusiastic about sports and one of the stars on the football team. You have talked to him about applying himself in class because you think he is smart enough to make it in college. You have told him that he could make better grades if he applied himself, but he would rather pursue his interests in sports and stock car racing. He is seldom prepared for class, but has never given you any trouble in terms of his class behavior. He had a C average on the four quizzes you gave, then got a D- on the six-week test. What grade you would assign Ronnie for the six-weeks? A, B, C, D, F

#8 (1,1,2,2,2)

Robert never wants to do any more than he absolutely has to in your class. He has gym just before your class and often has to hurry to get there on time, but he always makes it. However, he seldom has his homework done, always wanting more time to complete it. He offers excuses such as he had to go visit a sick relative or left his book at school, etc. He had a D+ average on the four quizzes you gave during the six-weeks and a C on the six-week test. You think he's much smarter than his grades indicate. Other teachers have told you that he does very well in their classes. What grade you would assign Robert for the six-weeks? A, B, C, D, F

#9 (1,2,1,1,1)

Henry, bothers you a lot by saying something "cute" at inappropriate times, for example, after the bell for dismissing class has rung and you are trying to give the next day's assignment. You have seen his permanent record, and know that he placed at the 41st percentile on the last standardized ability test he took. Often in class, even though he does his homework and asks questions, you just can't stand his "cute" attitude. He had a C average on the four quizzes you gave and got a D+ on the six-week test. What grade would you assign Henry for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd). Vignettes and Factor Level Combinations

#10 (1,2,1,1,2)

Hugh is a poor reader and his records show that in elementary school he was in Chapter I reading. He tries hard in class but because of his reading has difficulty. He also sometimes acts immature and gets into trouble. You had to write him up on different occasions for pulling hair and sticking people with a pin. His records show that he is at the 41st percentile in ability for students his age on a recently administered standardized ability test. He had a F average on four quizzes you gave during the grading period and a D+ on the six-week test. What grade you would assign Hugh for the six-weeks?

A, B, C, D, F

#11 (1,2,1,2,1)

If you could give all the students in your class the individual attention they need you are sure many would do much better. Jamie is one who seems to need individual attention. He is in your last period class. Jamie does his homework and if given individual attention does fine in class. However, left on his own he often starts playing around and gets into trouble. You had to write him up twice during the six weeks for horsing around in class and not stopping when you asked him repeatedly to stop. He can bother you so much in class that you often check the daily absentee roster to see if he came to school that day. You and some of his other teachers feel, from his class work and tests, that he has above average intelligence. However, during the last six-weeks he had a B- average on the four quizzes you gave and got a D- on the six-week test. What grade would assign Jamie for the six-weeks? A, B, C, D, F

#12 (1,2,1,2,2)

Eugene is the type of student other students like to pick on. You try to protect him from the other students in class by having him sit in the front row. You have spoken to him about his feelings, but he still has a lot of trouble with the teasing and often loses his temper. He always completes his homework assignments and contributes in discussions during class. He is a bright student, with a standardized ability test score at the 84th percentile. He got an A on your six-week test, but had only a C- average on the four quizzes you gave. What grade you would assign Eugene for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd). Vignettes and Factor Level Combinations

#13 (1,2,2,1,1)

Donnie does a lot of little extra things for you, like running errands to the office and helping straighten up the desks after class. You and Donnie get along very well. Based on his classwork and the opinion of other teachers you judge he would be placed in the low ability range. However, even though his ability is on the low side, he works hard in class and seems genuinely interested. He even asks for extra work to do outside of class. On the four quizzes you gave during the six-weeks he had a B average, but got a C- on your six-week test. What grade you would assign Donnie for the six-weeks? A, B, C, D, F

#14 (1,2,2,1,2)

Mike is from a family of hard workers. You had his sister last year and she was an "A" student. Just like his sister, he works hard too. He always has his homework done which is always extremely neat. He is trustworthy and someone you can leave "in charge" if you have to go out. From his permanent record you know he is at the 46th percentile in ability, a little below average. He got an A on your six-week test, but had a C- quiz average on the four quizzes you gave. What grade you would assign Mike for the six-weeks? A, B, C, D, F

#15 (1,2,2,2,1)

You have only had a few students like Steven, you wish there were more. He is in your first period class. He wants to go to a top notch college and always wants talk to you about where he should go to college and what he should major in. Steve is always well prepared for class. He is also considerate of you and the other students. In many ways he is a model student. You feel he is bright, but not extremely so. His standardized ability score places him at the 82nd percentile. He had an B+ average on the quizzes you gave, but only got a D+ on your six-week test. You asked him what happened on the test and he said he spent too much time working on a project for another class and didn't study like he should have. What grade would you assign Steven for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd). Vignettes and Factor Level Combinations

#16 (1,2,2,2,2)

Nick is in your class. He has surprised you on several occasions by picking up concepts and ideas ahead of any of the other students in class. You feel that he is one of your brighter students. He always has his homework done and often asks if there is anything else he can do for extra credit. He often helps you by doing extra things like calling role. During the last six-weeks you gave four quizzes. He didn't study for one of the quizzes you gave, said he thought he knew the material, and made a terrible grade. However, he still had a B- average on the quizzes. He then got an A on his six-week test. What grade would you assign Nick for the six-weeks? A, B, C, D, F

#17 (2,1,1,1,1)

You learned only recently that one of your students, Helen, was pregnant. Yesterday, she had a little girl. She is supposed to soon return to class. Hopefully, when she returns she will have a more positive attitude toward school. Before she left, she would never participate in class. One day you had to write her up for talking and then cursing under her breath when you told her to be quiet. Her grades during the last six-weeks were a C- average on the four quizzes you gave and an F on her six-week test. Her cumulative folder shows that on the last standardized ability test she took that she was at the 38th percentile. What grade would you assign Helen for the six-weeks? A, B, C, D, F

#18 (2,1,1,1,2)

Sandy does not like you and has told you so in class. While it embarrassed you when she said it, you were angry because she caught you so off-guard. However, you think of her as a sad case. She stays "in trouble" at school and has been suspended before. You have talked to other teachers about her and they all agree that she has a bad attitude, especially toward any authority figure. It's not surprising to you that she scored at the 39th percentile on the last standardized ability test she took. However, she had a C- on the quizzes you gave (you think perhaps out of spite) and got a C on her six-weeks test. What grade would you assign Sandy for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd). Vignettes and Factor Level Combinations

#19 (2,1,1,2,1)

Janet comes to class with a chip on her shoulder. She lives with her mother, three brothers and two sisters. As the oldest of six children, she tells you she has to baby-sit all the time and that is why she never has her homework completed. She misses at least one day of school per week on average. While these usually fall on Mondays or Fridays she managed to take all of the quizzes you gave during the last grading period. Her permanent record shows her at the 80th percentile in ability on the most recent standardized test she took. Her quiz average for the six-weeks was a B-, but she failed your six-week test. What grade would you assign Janet for the six-weeks? A, B, C, D, F

#20 (2,1,1,2,2)

Sometimes there are students you just don't understand and Judy is one of them. Her other teachers and yourself, from her classwork and quizzes and tests feel that she is very bright and could go to college if she applied herself. However, she "hangs out" with students who are always getting in trouble. Judy shows little interest in your class and often acts immature - giggling and passing notes. You have "written her up" for talking. She seldom has her homework done. During the past six-weeks she had a D- average on the four quizzes you gave and a D on your six-week test. What grade would you assign Judy for the six-weeks? A, B, C, D, F

#21 (2,1,2,1,1)

During this six-weeks you have been asking your students to turn-in answers to certain questions at the end of each chapter in the textbook. The work is not graded but makes the students read the chapter. Susan is a very shy person and seldom completes these assignments. She is never a problem in class, but you are concerned about her being too quiet. Her permanent record shows that she scored at the 42nd percentile on the last administered standardized ability test. She had a B- average on the quizzes you gave during the six-weeks. However, she got a D- on her six-weeks test. What grade would you assign Susan for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd). Vignettes and Factor Level Combinations

#22 (2,1,2,1,2)

Terry had a D+ on the four quizzes you gave during the six-weeks and a C on your six-week's test. She always follows your class rules and even helps you enforce them. You and some of her other teachers feel Terry has somewhat low ability. Based on her past test performance and classwork. She is such a nice girl, you wish she could transfer some of her good behavior into more effort and enthusiasm for class. You know that her father has not worked for over a year and that she helps her mother clean houses to make ends meet. She has told you that the reason she never has her homework done, or has read what you assign for class, is because she was working. What grade would you assign Terry for the six-weeks? A, B, C, D, F

#23 (2,1,2,2,1)

Carol doesn't want to do very much in your class. However, you know that she is considered a smart student by many teachers. Therefore, you think that either she doesn't like your class or maybe you. She is polite to you and everyone in class, always raises her hand when she wants to speak, etc., but "sluffs-off" on her classwork whenever she can. Recently, she got a D on your six-weeks test and had a C+ average on the four quizzes you gave in class during the grading period. What grade would you assign Carol for the six-weeks? A, B, C, D, F

#24 (2,1,2,2,2)

Patty seems to have little regard for her future. You feel that she could be anything she wanted to be, and yet she puts all of her effort into the "dating game". You have talked with her concerning her priorities and about her lack of interest in class. But, she remains unconcerned. There is no doubt in your mind that she is bright and should go on to college. However, listening to you and applying herself in class seems to be the last thing on her mind. While she is considerate of her classmates, in terms of her classwork she does very little. It was a surprise to you when she had a B- average on the four quizzes you gave and then got an A on the six-week test. What grade would you assign Patty for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd).

Vignettes and Factor Level Combinations

#25 (2,2,1,1,1)

Janice, a student in your second period class, denotes an hour every morning on her hair and makeup. Her cumulative folder shows that she is at the 39th percentile in ability on the last standardized ability test she took. She often manages to get into some sort of trouble. Recently, when she asked your permission to go to the restroom, another teacher caught her using the pay phone in the cafeteria and wrote her up. This was not the first time she had abused your trust. You know that her parents want her to go to college. Janice has come to you and asked if there is anything she can do to improve her grade. However, her grades dropped during the six-weeks from an A-average, on the four quizzes you gave, to a C- on your six-week test. What grade would you assign Janice for the six-weeks?
A, B, C, D, F

#26 (2,2,1,1,2)

Sally is a girl in your class that is usually prepared for class. However, you have moved her to keep her out of mischief. This six-weeks she has been sitting in the middle of the front row, so you can keep an eye on her. Last grading period, you caught her and another girl cutting class and wrote them both up. Sally's cumulative record shows that she is at the 43rd percentile in ability on the most recent standardized test she took. Her new seating arrangement must be working because during the last six-weeks she had a C+ average on the four quizzes you gave and a B on your six-weeks test. What grade would you assign Sally for the six-weeks? A, B, C, D, F

#27 (2,2,1,2,1)

You had to do something you really dislike doing several months ago in your first period class. You had to write a student up for cheating. You caught Crystal, a student of yours, cheating on one of your tests. It really surprised you, because she is bright and participates in class. She had notes hidden under her test paper. You saw her looking at them and caught her red-handed. While you are still a little suspicious of her, you are now in another grading period. Crystal had a B+ on the four quizzes you gave and a D+ on the six-weeks test. What grade you would assign Crystal for the six-weeks?
A, B, C, D, F

Figure B-1 (cont'd).

Vignettes and Factor Level Combinations

#28 (2,2,1,2,2)

Thelma is absent from school a lot and often tardy to your class when in attendance. She missed eight days the last six-weeks. It's bad enough that she is tardy but she also makes a lot of noise when she comes in. When you have spoken to her about her tardiness and her disrupting class, her attitude has always been "so what". You've seen her permanent record and know that she placed in the top quartile on a standardized ability measure. She is a talker and likes to take part in the discussions you have in class. During the last six-weeks she had a C- average on the four quizzes you gave (she was in school for all of them) and had an A- on her six-weeks test. What grade you would assign Thelma for the six-weeks? A, B, C, D, F

#29 (2,2,2,1,1)

You have enjoyed attending the girls basketball games this year and watching Denise develop into a fine athlete. Although she puts forth the same level of effort in your class, her academic ability lags far behind her athletic ability. Her reading comprehension is very limiting to her, and as a consequence she has difficulty participating meaningfully in class discussions. You leave her "in charge" of the class whenever you have to go out of the room because you trust her. During the last grading period she had a B- average on the quizzes given and a C on the six-weeks test. What grade you would assign Denise for the six-weeks? A, B, C, D, F

#30 (2,2,2,1,2)

One of your students, Bobbie, wants to be a nurse. She has talked with you about how she plans to take classes at a nearby college. She had a D on the quizzes you gave, but got a B- on your six-weeks test. You and Bobbie's other teachers feel that from her classwork, tests, and observations that she has limited ability, and will have to work very hard to be successful in college. She always has her homework done in your class and likes to help you and the other students in the class. What grade would you assign Bobbie for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd). Vignettes and Factor Level Combinations

#31 (2,2,2,2,1)

Stacie is in your third period class. She is a cheerleader, president of her class, and usually an honor role student. She will probably be the homecoming queen because she is so attractive. She is more than a little snobby, however. She is always considerate of you and the other students. During the last grading period in your class her grades went down. She had an A on the quizzes you gave, but then got a C- on your six-weeks test. What grade you would assign Stacie for the six-weeks?

A, B, C, D, F

#32 (2,2,2,2,2)

Kathy could really benefit from college, however she probably will not go, but rather start working at a nearby factory when she graduates. It's a shame because she is studious, hard working, talented, well-mannered, and genuinely interested in your class. She had a B average on your quizzes and an A- on the six-weeks test. What grade you would assign Kathy for the six-weeks? A, B, C, D, F

Figure B-1 (cont'd).

Vignettes and Factor Level Combinations

Table B-1

Vignette Factor Level Combinations, Letter Grade Changes, and Numeric and Letter Grade Averages by Grade Change

<u>Vignette/Levels</u>	<u>Quiz Avg./Test Grade</u>		<u>Num./Letter Average</u>		<u>Num./Letter Change</u>	
#2 (1,1,1,1,2)	D	D+	1.15	D	+0.3	(1/2)
#4 (1,1,1,2,2)	C	C+	2.15	C	+0.3	(1/2)
#18 (2,1,1,1,2)	C-	C	1.85	C	+0.3	(1/2)
#20 (2,1,1,2,2)	D-	D	0.85	D	+0.3	(1/2)
#8 (1,1,2,2,2)	D+	C	1.65	C	+0.7	(1)
#9 (1,2,1,1,1)	C	D+	1.65	C	-0.7	(1)
#22 (2,1,2,1,2)	D+	C	1.65	C	+0.7	(1)
#26 (2,2,1,1,2)	C+	B	2.65	B	+0.7	(1)
#29 (2,2,2,1,1)	B-	C	2.35	C	-0.7	(1)
#32 (2,2,2,2,2)	B	A-	3.35	B	+0.7	(1)
#7 (1,1,2,2,1)	C	D-	1.35	D	-1.3	(1 1/2)
#10 (1,2,1,1,2)	F	D+	0.65	D	+1.3	(1 1/2)
#13 (1,2,2,1,1)	B	C-	2.35	C	-1.3	(1 1/2)
#16 (1,2,2,2,2)	B-	A	3.35	B	+1.3	(1 1/2)
#23 (2,1,2,2,1)	C+	D	1.65	C	-1.3	(1 1/2)
#24 (2,1,2,2,2)	B-	A	3.35	B	+1.3	(1 1/2)
#17 (2,1,1,1,1)	C-	F	0.85	D	-1.7	(1 1/2)
#1 (1,1,1,1,1)	C-	F	0.85	D	-1.7	(1 1/2)
#5 (1,1,2,1,1)	C-	F	0.85	D	-1.7	(1 1/2)
#30 (2,2,2,1,2)	D	B-	1.85	C	+1.7	(1 1/2)
#6 (1,1,2,1,2)	C-	A-	2.70	B	+2	(2)
#11 (1,2,1,2,1)	B-	D-	1.70	C	-2	(2)
#15 (1,2,2,2,1)	B+	D+	2.30	C	-2	(2)
#25 (2,2,1,1,1)	A-	C-	2.70	B	-2	(2)
#27 (2,2,1,2,1)	B+	D+	2.30	C	-2	(2)
#28 (2,2,1,2,2)	C-	A-	2.70	B	+2	(2)
#21 (2,1,2,1,1)	B-	D-	1.70	C	-2	(2)
#12 (1,2,1,2,2)	C-	A	2.85	B	+2.3	(2)
#14 (1,2,2,1,2)	C-	A	2.85	B	+2.3	(2)
#3 (1,1,1,2,1)	B	D-	1.85	C	-2.3	(2)
#31 (2,2,2,2,1)	A	C-	2.85	B	-2.3	(2)
#19 (2,1,1,2,1)	B-	F	1.35	D	-2.7	(2 1/2)

Note. Factor order and coding of vignettes are: sex (male=1, female=2), effort (low=1, high=2), behavior (unacceptable=1, desirable=2), ability (low=1, high=2), and grade change (decrease=1, increase=2).

Ordering of Information is by Numeric and Letter Grade Changes.

Table B-2

Presentation Order for Vignettes in Forms 1, 2, and 3

<u>Original Number</u>	<u>Form 1</u>	<u>Form 2</u>	<u>Form 2</u>
1	7	31	15
2	32	20	8
3	14	22	2
4	17	5	17
5	20	8	20
6	27	11	10
7	28	32	12
8	1	13	22
9	11	3	24
10	26	10	9
11	3	15	31
12	29	17	5
13	24	28	28
14	5	19	13
15	15	23	3
16	22	26	26
17	19	7	19
18	13	21	1
19	30	18	6
20	6	30	14
21	18	6	18
22	8	12	16
23	4	16	32
24	16	24	4
25	23	27	27
26	10	2	23
27	2	14	30
28	25	9	11
29	31	29	7
30	12	4	29
31	9	1	21
32	21	25	25

AN EXERCISE IN GRADING PRACTICES

Directions:

You are asked to read each of thirty-two vignettes. In each vignette a fictitious student is described. After reading a vignette, you are asked to assign a grade to the hypothetical student depicted. Please respond to the vignettes in terms of one of the academic classes you are actually teaching this term. Describe the particular class you have in mind below.

Subject _____ Grade level _____

After reading each vignette assign the grade you would award to the hypothetical student by circling one of the letter grades at the end of the vignette. For the purpose of this exercise assume that school policy permits recording only letter grades on report cards. Accordingly, refrain from adding + or - even though you may be tempted. Assume that your grading policy calls for the quiz grades and test grades to be given equal weight.

It is essential to the analysis that you record a grade for every vignette. Consequently, you are asked to assign a grade to each vignette even if you may be uncomfortable in doing so in some cases. You are invited to write comments regarding any of the vignettes or problems you experience doing this exercise on the last page.

Would you please provide the following information with respect to yourself and teaching.

1. How many years teaching experience do you have, counting this year? _____
2. What is your highest level of educational attainment?
(a) BS/BA Degree ____ (c) Doctoral ____
(b) Masters Degree ____
3. What is your gender? Male ____ Female ____
4. What is your ethnicity?
(a) Am. Indian and Native Alaskan ____ (d) Hispanic ____
(b) Asian and Pacific Islander ____ (e) White ____
(c) Black/Afro-American ____ (f) Other ____
5. What subjects and grade levels are you teaching this year?
(a) _____
(b) _____
(c) _____
(d) _____
(e) _____

Figure B-2. Directions and Questions Attached to Each Packet of Vignettes

APPENDIX C

Review of Selected Literature

POLICY CAPTURING METHODOLOGY

Policy capturing is a methodology designed to assess the importance people attach to factors considered when reaching a decision. For example, policy capturing could be used to assess the relative importance people attach to such factors as styling, economy, fuel efficiency, etc. when purchasing a new car. Most policy capturing studies require a panel of judges to read a series of vignettes and make a decision based on the information contained in each vignette. Each vignette differs systematically from the others along one or more dimensions, or informational sources.

The methodology has been applied in different studies as indicated by the studies cited in Table C-1. However, there is much diversity concerning the characteristics of studies using a policy capturing methodology. The characteristics of each of the cited studies are contained within the table's columns. For the purpose of application of the methodology, one of the cited studies will be described in detail below.

Zedeck and Kafry (1977) used policy capturing methodology in a study to evaluate criteria associated with nurses' job effectiveness. Sixty-seven nurses were asked to rate 40 hypothetical profiles or vignettes depicting 40 different nurses. The nurses were asked to rate each of the vignettes of hypothetical nurses on a 7-point scale for overall effectiveness, where '1' was

very ineffective and '7' was very effective. The 40 vignettes were created by varying combinations of behavioral descriptions on one of three performance levels (poor, average, good) for nine different informational elements: (1) organizational ability; (2) assessment of family needs and problems; (3) interviewing skills; (4) planning, decision-making, and recommending; (5) knowledge; (6) follow-up actions; (7) written communications, (8) system orientation; and (9) professional growth.

The 67 subjects in the study were each presented 40 vignettes, far fewer than all possible combinations (19,683) of the nine dimensions with three levels of performance per dimension. The researchers indicated that a major concern was in creating vignettes which represented realistic profiles of nurses, rather than chance combinations of the attributes under investigation. Each of the vignettes selected had to satisfy the following criteria: (1) consensus of four industrial/organizational psychologists regarding which behavioral combinations were most realistic views of nurses roles; (2) informational elements being approximately normally distributed; and (3) no intercorrelations among informational elements.

Sixty-seven separate regression equations were created, one for each nurse serving as a judge. The nine elements in each vignette served as an independent variable and the effectiveness rating of the nurses served as the dependent variable. The resulting regression coefficients for each nurse's

regression equation were then interpreted as the amount of importance placed by a nurse on the nine elements in evaluating each vignette.

The 67 participating nurses were divided into two groups, 35 Public Health Nurses (PHN) and 32 Hospital Registered Nurses (RN). Comparisons of the two groups revealed differences in importance placed on the nine elements. For the PHN's, the four most important elements were: (1) organizational ability; (2) assessment of family needs and problems; (3) system orientation; and (4) planning, decision making and recommending. By contrast, the RN's placed greatest importance on: (1) empathy; (2) clinical knowledge; (3) professional growth; and (4) leadership ability.

Concerns of Policy Capturing Methodology

Policy capturing methodology has been used in studies concerning school quality, faculty evaluation, and student retention (Anderson, 1977; Borko & Cadwell, 1982; Hobson, Mendel, & Gibson, 1981; Jaeger & Usher, 1992; Tomchin, 1989); management and executive performance (Stahl & Zimmerer, 1984; Stumpf & London, 1981); military performance (Taylor & Wilsted, 1974); and, nurse effectiveness (Zedeck & Kafry, 1977). The major practical concerns for studies using policy capturing methodology include (1) the number of informational elements or factors and corresponding levels, (2) number of vignettes generated, and (3) the relationships among the informational elements. Differences across these characteristics found in the above cited studies are listed in the last three columns of Table C-1.

While evidence is that policy capturing methodology shows much diversity in use, it is not without criticism. Concerns include: (1) construct validity of profile simulations; (2) methodological issues involving rater profiles; (3) objective and subjective estimate comparisons; and (4) systematic differences among raters. Each of these topics is discussed below.

1. Construct Validity of Profile Simulations

Hobson and Gibson (1983) question whether or not "captured rating policies are construct valid representations of 'true' rating policies" (p.640). They contend that because most depictions of situations (vignettes) used in studies are artificial, in that, profiles are usually constructed by researchers to specifically include certain levels of criterion elements, profiles can easily differ "drastically from actual settings - thus calling into question the construct validity of the resultant captured policies" (p.640).

Additional deficiencies in using simulated profiles include: (1) not providing enough descriptors within profiles for a judge to formulate an opinion of the situation; (2) presentation of information cues to raters which are too general in nature; and (3) profiles often lacking "realness" (Hobson & Gibson, 1983). However, counter-arguments to some of these claims are presented by Jaeger and Usher (1992), pointing out that: (1) distractions of more descriptors within profiles are problematic, in that profiles become too lengthy, (2) if cues are more specific, descriptions become too narrow and

might cause judges to become too focused on specific attributes rather than viewing the profile in its entirety, and (3) familiarity with persons judged might result in a halo effect, undermining the construct validity of any ratings.

2. Methodological Issues Involving Ratee Profiles

Hypothetical Versus Real Ratee Profiles. Of the policy capturing studies cited in Table C-1 only the study by Taylor and Wilsted (1974) was conducted with real, as opposed to simulated profiles. Their study utilized Air Force officers performance profiles from personnel records. Such real profiles are referred to in the literature as "field studies" and are considered by Hobson and Gibson (1983) as the "ideal" in profile presentation. The hypothetical profile, or vignette, is used in most policy capturing studies because of its utility in combining different pieces of information which the researcher would like to evaluate. However, it is beneficial to pilot test any hypothetical vignettes. Vignettes in the present study, involving teachers' grading practices, were reviewed by teachers in a pilot study to determine how "believable" they were.

Ratee Performance Format. Formats for depicting behavior in profiles range from a simple numerical presentation to a paragraph description (Zedeck & Kafry, 1977) to a graphic presentation with behaviorally defined anchor points (Hobson et al., 1981). Differences have been found among different profile presentations. Anderson (1977) presented evidence that

numerically defined ratee profiles are evaluated more consistently than verbal ones in paragraph form.

A major practical concern in using policy capturing methodology is eliciting the cooperation of participants to spend a rather large amount of time reading, reviewing, and responding to many vignettes containing fictitious depictions of information. The amount of time estimated by the participant in completing an exercise involving, often, 32, 40, or 80 vignettes is of special concern for the researcher. Lengthy profile presentation formats can easily take up an entire page, meaning that a subject would be handed 32 or more pages to review. If such an approach is taken, the subject, if given the opportunity, may very well decline to participate.

A much more manageable format was included in the present study of teachers' grading practices. Brief narratives combining student characteristics were created which were no longer than a "short" paragraph. Usually, three vignettes could be placed on a page. While the present study created 32 vignettes for participants to review, three different vignette orderings, or forms, were also created. This was done to minimize comparison and possible copying of responses by participants. Each participant received a packet including directions and one of the three vignette orderings. Two of the packets were nine pages in length. The third was ten pages.

3. Objective Versus Subjective Estimates

In attempting to show that differences occur between what people think and what they actually do, several studies using policy capturing methodology have obtained both objective estimates (derived from regression coefficients from individual's regression equations) and subjective estimates from subjects (e.g., Hobson et al., 1981; Stumpf & London, 1981; Zedeck & Kafray's, 1977). For example, Zedeck and Kafray (1977) gathered both subjective and objective estimates from subjects in their study. Nurses were asked, after they had completed rating the nurses in the 40 hypothetical profiles, to distribute 100 points to the nine informational elements according to each element's relative importance. When the relative versus subjective weightings of the elements were evaluated, it was found "that fewer elements are used in over-all assessment than when subjective, explicit weights are used to explain decision variance" (p. 281).

The smaller number of factors explaining decision-making variance in Zedeck and Kafray's study is mirrored in other policy capturing studies where both objective and subjective ratings were obtained (e.g. Hobson et al., 1981; Stumpf & London, 1981). This supports the contention that judges use fewer informal cues in decision-making than is usually thought (Hobson & Gibson, 1983). The present study chose not to collect subjective information from teachers due to the possibility that, in doing so, the systematic presentation of informational cues within the vignettes would be revealed.

4. Systematic Differences Among Raters

Various statistical clustering procedures have been used by researchers to group raters whose rating policies are similar (Hobson et al., 1981, Stump & London, 1981; Zedeck & Kafry, 1977). For example, Zedeck and Kafry used a judgment analysis (JAN) procedure in analyzing similarities among the group of 35 Public Health Nurses' (PHN) relative weights for the nine factors. When the JAN procedure was employed, two clusters emerged, one, placing greater emphasis on the elements "planning, decision-making, and recommending" while the other cluster placed greater emphasis on two elements, "knowledge" and "written communications". While the clustering technique seemed like a good idea to the authors, it was discovered that, even though differences existed between clusters, the elements identified were not those receiving the highest relative values from the subjects. Rather, the group of PHN's considered the elements "organizational ability", "system orientation", and "assessment of family needs and problems" most important. This was evident in both clusters. In conclusion, the researchers felt that, due to the similarity of relative means for each of the informational elements in each of the clusters, the decision-making processes of the PHN raters were basically homogeneous.

USE OF FACTORIAL DESIGNS IN POLICY CAPTURING STUDIES

1. Concerns of the Present Study

A problem inherent to studies using policy capturing methodology is interpretation of the effects of the informational elements or factors. If the factors within a policy capturing study are correlated, interpretation of a factor's effect is confounded with the effects of other factors. Such concern was considered especially important in the present study since the objective was in evaluating the influence, independently, of non-achievement factors on teachers' grading practices.

While concern for factor interpretation was important in the present study's design, several other concerns needed also to be considered. First, it was considered important to present all possible combinations of the factors under consideration to subjects, wherein the levels of each factor could be combined with each level of the other factors. Second, it was felt that it was necessary to keep the number of vignettes manageable in number so that teachers would be able to respond to them during teachers' meetings, where they were most often contacted. Third, directly tied to the number of vignettes created from all possible combinations of factor levels was the ratio of vignettes to factors. This ratio needed to be kept small, ideally 10 to 1.

2. Factorial Design

In addressing the concerns cited above, it was decided that a complete factorial design would be used. Such a design yields zero correlations among

factors and would permit unambiguous interpretation of the effect of each factor on the criterion judgment. In addition, a completely crossed factorial design would create all possible combinations of the factor levels which would be presented to subjects. When regression analysis is applied to data from a factorial design the regression weights are independent of one another.

Number of Possible Combinations

Much concern in policy capturing studies is directed toward creating believable and realistic vignettes. However, this concern is compounded by the additional problem of the number of vignettes created easily becoming unrealistic in terms of the amount of time needed for their review by participants. For example, suppose a study would like to have its participants consider all combinations of five factors, each factor having three levels of performance. If the participants considered vignettes containing all the possible combinations, there would be 243 vignettes to consider. If there were six factors with two levels of performance, there would be 64 vignettes. If there were three factors with four levels of performance, there would be 81 vignettes.

3. The Multiple Regression Model

Hoffman (1960), one of the ground-breakers of policy capturing methodology, pointed out that part of the value of the policy capturing methodology was that it combined information used in decision-making in such a way that importance was sometimes placed on information which

otherwise would go undetected. He suggested that the linear multiple regression model could be used "paramorphically" in representing human judgement. Oskamp (1967) clarified what Hoffman meant by the term "paramorphical representation". He stated, "a high multiple R means that a judge's decisions may be 'reproduced' by a linear, additive combination of predictor variables; however, it does not mean that the judge himself used a linear, additive method of combining variables" (p.414).

The Additive Multiple Regression Model

While much of the research in the area of decision-making behavior has focused on the multiple regression model, in almost every instance, where the relationship between individual's decisions and informational elements or factors were examined, the additive linear regression model has adequately modeled the relationship (Jaeger & Usher, 1991; Stahl & Zimmerer, 1984). Elimination of interaction terms in multiple linear regression models is seen throughout behavioral decision theory literature when policy capturing methodology has been employed (e.g., Hoffman, 1960; Jaeger & Usher, 1991; Laughlin, 1978; Stahl & Zimmerer, 1984). As Jaeger and Usher note the additive model is recommended because, first, reasonably high coefficients of determination have been observed in most policy capturing studies and, second, only extremely small increments "in proportions of predicted variation of criterion scores" (p. 14) are shown when interaction terms were included.

Ratio Between Number of Vignettes and Factors

As noted above for the present study, concern was directed to the ratio between the number of vignettes and number of factors ideally being 10 to 1. Several researchers recommend, when using multiple regression analysis, setting a minimum standard ratio (e.g, Hobson & Gibson, 1983; Nunnally, 1978; Pedhazur, 1982). However, in studies using policy capturing methodology, much importance is, usually, placed on creating believable and meaningful hypothetical vignette simulations. For example, in deciding on which vignettes to use in their study, Zedack and Kafry elicited consensus of four industrial/organizational psychologists regarding which behavioral combinations were most realistic views of nurses roles.

In designing the present study, the believability of vignettes along with creating a manageable number of profiles to be presented to teachers were major concerns, as was there not being any correlation among factors. In addition, it was considered important to include all possible combinations of the five informational cues (each having two levels) within vignettes. In so doing, the resulting ratio between vignettes (32) and independent variables (5) was slightly above 6:1.

4. Fractional Factorial Design

It should be pointed out that another approach used in policy capturing studies to maintain a high ratio between the number of vignettes and informational elements has been to use a fractional factorial design, splitting

descriptions of informational elements into blocks (e.g., Borko & Cadweld, 1982; Stahl & Zimmerer, 1984). While this reduces the number of vignettes subjects are asked to review, difficulty arises when the weights of individual subject's regression equations are interpreted.

A fractional factorial design was not used in the present study for two reasons. First, it was considered important to present all possible informational element combinations of the five factors in the vignettes to each subject. This meant constructing 32 vignettes, which seemed a manageable number for participating teachers to respond to. Second, if each subject does not consider all combinations of informational cues, i.e., review all 32 vignettes, the relative weight a factor may receive would be a potentially confounded estimate of the importance placed on that factor by the subject.

SUMMARY

Some general conclusions concerning policy capturing methodology drawn from this review, include, first, that there is support for the methodology, evidenced by the diversity and increased use in recent years. Second, utilizing a factorial design is considered desirable in policy capturing studies because it allows for the unique effects of each of the factors within a study to be interpreted as indices of importance within an individual's decision-making.

Third, the general additive linear multiple regression model is the analytic procedure of choice in most policy capturing studies and seems to work well in describing raters' decision-making. Most of the studies cited in this review used the additive model and reported consistently high rater R-squares.

Fourth, several policy capturing studies have employed clustering procedures in searching for similarities among raters. Generally, studies have found that a single rating policy could not represent the decision-making behavior of raters. Rather, raters seem to make decisions uniquely, varying the amount of importance they place on certain factors.

Finally, several policy capturing studies have elicited both objective (relative) and subjective ratings from participants. In so doing, most have found that subjects, generally, had poor insight into their own decision-making, thinking they were placing importance on certain factors while actually placing importance on "other" factors.

Table C-1 Summary of policy-capturing literature in terms of study's characteristics

Reference	Judge	Dependent Variable	Profile Type	# Informational Cues	# Profiles	Intercorrelations
Anderson (1977)	164 high school teachers	teacher quality	Narrative & numerical description	4, 6, 8	36	Orthogonal
Borko & Cadwell (1982)	41 elem. teachers	Judgements aptitudes; Preinstruct. decisions (5-pt. & 6-pt. scales)	Narrative description	6 (2 levels each)	32, one-half replicate of 64	Orthogonal
Feldman & Arnold (1978)	62 graduate students	Willingness to accept position (continuous scale)	Narrative descriptions	6 (2 levels each)	64	Orthogonal
Hobson, Mendel, & Gibson (1981)	20 faculty members	Performance (9-pt. scale)	Line graph w/ numerical ratings	14 (3 levels each)	100	Orthogonal
Jaeger & Usher (1991)	28 educ. leaders	Type of School (choice of 1 of 3 categories)	Graphical Format	8 (3 levels each)	80	Logically derived by experts
Stahl & Zimmerer (1984)	42 executives	Candidate firms (rating of whether or not to acquire firm)	Narrative descriptions w/ categorical	6 (2 levels each)	32 (one-half replicate of 64)	Orthogonal

(Table C-1., continued)

Reference	Judge	Dependent Variable	Profile Type	# Informational Cues	# Profiles	Intercorrelations
Stumpf & London (1981)	43 managers	Performance rating (5-pt. scale)	Narrative description with numerical ratings	5, plus 26 interaction terms	48	Orthogonal
Taylor & Wilsted (1974)	25 Air Force NCO's	Performance rating	Field Setting (no simulations)	10 (7 levels each)	25	Empirically determined (post hoc)
Tomchin (1989)	98 elem. teachers	Retention decisions (promote/retain)	Narrative description w/ graphics	6 (2 levels each)	40	Orthogonal
Zedeck & Kafry (1977)	67 nurses	Effectiveness rating (7-pt. scale)	Narrative descriptions	9 (3 levels each)	40	Orthogonal

APPENDIX D

TEACHERS' COMMENTS

Tch-1 Subject area/Grade: Geometry, Algebra I, 9-12
Experience: 22 years
Gender: Female
Discrepancies: 6
Significant factors: Grade Change

"Behavior problems like effort and good/poor standardized test scores should not be a factor in giving a grade."

"When a student's average is on the top of one letter scale, then effort and percentile (score) will be a factor for me - I usually go with the next grade."

"I believe numeric grades should always be recorded."

Tch-3 Subject area: Chemistry, 11
Experience: 4 years
Gender: Female
Discrepancies: 5
Significant factors: none

"The extra information provided - in addition to the grades achieved by these students - would be helpful in choosing day-to-day teaching techniques; however, it rarely made a difference when I assigned six-weeks grades."

Tch-4 Subject area/Grade: Geometry, 10
Experience: 4 years
Gender: Female
Discrepancies: 5
Significant factors: Ability and Grade Change

"I consider 'circumstances' when the grade is less than half a point from the next grade level"

Tch-10	Subject area/Grade:	English, 11
	Experience:	33 years
	Gender:	Female
	Discrepancies:	13
	Significant factors:	Grade Change

"I am disturbed that your grading criteria only includes test scores and no daily grades. I am sure that it was meant to keep this as simple as possible, but testing situations are often traumatic for students. Not good criteria. It is obvious that you wish to note whether or not there is correlation between a student's behavior and his grade (How much does student behavior influence teachers?). I don't feel this is structured to show a strong or weak influence."

Tch-11	Subject area/Grade:	Math, 9-12
	Experience:	31 years
	Gender:	Male
	Discrepancies:	13
	Significant factors:	Grade Change

"Academics, attendance, discipline, and other should be dealt with on their own as far as grades are concerned. Although grades and discipline, attendance, personality traits, home life et al. are closely interrelated, grades should be a measure of academic achievement only."

"After an excess of 30 years of teaching and administration (15), I am convinced that we need to measure academia with academic methods and treat these other things as they need to be treated, etc."

Tch-13	Subject area/Grade:	English, 12
	Experience:	20 years
	Gender:	Female
	Discrepancies:	0

"Grades are very objective - circumstances don't really make any difference unless the grades are .5 or above."

"All my answers were based on a 10 pt. scale, where 'F's' are concerned - it really depends on how low the F was."

"I think there needs to be more grades than were presented in the vignettes"

"Standardized scores should have no bearing on a six-weeks grade! They are questionable and unreliable!"

"Grades should be based on various assessments - not just quizzes and tests - portfolio assessments needs to be used with conferences between student's and teachers to arrive at a 'grade'."

Tch-14	Subject area/Grade:	English, 11
	Experience:	23 years
	Gender:	Female
	Discrepancies:	4

"I sincerely hope that you are not planning to draw the obvious conclusions from your tabulated responses. These vignettes lead me to believe that you perhaps plan to draw unjustified and unjustifiable conclusions when, in reality, the vignette presented assorted statements that do not happen - at least in our school"

Tch-15	Subject area/Grade:	Biology, 10-12
	Experience:	6 years
	Gender:	Female
	Discrepancies:	6
	Significant factors:	Effort

"Behavior should not be a factor in determining a student's grade."

"The grade should reflect the student's academic achievement."

"If the grade is borderline, the amount of sincere effort a student makes can influence the final grade."

Tch-17	Subject area/Grade:	English, 11 and 12
	Experience:	19 years
	Gender:	Female
	Discrepancies:	0

"Grades are based on student achievement. Behavior problems, personal problems, and individual student/teacher relationships

should not be used as a guide for determining achievement. My grade selections were based on my guess of an average. No other variables were or should be considered."

Tch-20	Subject area/Grade:	Chemistry, 11
	Experience:	23 years
	Gender:	Female
	Discrepancies:	3

"I have to admit that I looked at the quiz average and test average and mentally averaged them to get a grade. After a while, I quit reading your scenarios feeling they were irrelevant".

"Something you need to consider is the inconsistency of the + and -'s on your quizzes and tests, but not in the grade given at the end."

"One thing you mention in your scenarios is homework, doing it or not, but it appears that you do not 'grade' it. I think that it should be considered and graded in some way."

Tch-21	Subject area/Grade:	U.S. Government, 12
	Experience:	9 years
	Gender:	Female
	Discrepancies:	1

"As a 12th grade teacher, I try to be more 'results' oriented than 'cause' oriented. To a degree the vignettes are meaningless to me. My grade would be basically the average (a student) objectively earned."

Tch-22	Subject area/Grade:	History, 10 and 11
	Experience:	13 years
	Gender:	Female
	Discrepancies:	6
	Significant factors:	Behavior

"None of the circumstances would affect the student's grade in my class. My students' grades are determined by numerical average only."

Tch-23	Subject area/Grade:	English, 10-12
	Experience:	20 years
	Gender:	Female
	Discrepancies:	6
	Significant factors:	Grade Change

"Not enough information given to determine grades."

Tch-27	Subject area/Grade:	English, 10
	Experience:	23 years
	Gender:	Female
	Discrepancies:	6
	Significant factors:	Grade Change

"The information provided about each student is interesting and helpful in deciding how best to 'reach' a student, but academic grades are determined numerically. We have separate categories for effort and conduct grades."

Tch-28	Subject area/Grade:	Science, 9
	Experience:	30 years
	Gender:	Male
	Discrepancies:	12
	Significant factors:	none

"Typically, each grade is increased a letter from those indicated so there is justification for passing marks for those who tried at times to understand, but had failing averages. This keeps me from looking bad from a lot of failures and separates those students from a high percentage of students who do nothing except create disturbances. It also gets me out of some angry confrontations in which parents feel the child should have received a better grade. At the end of the year, most of those who do nothing except disturb (class) will also receive passing marks because the idea of saving taxpayers money outweighs any perceived gains from example or repetition, and saves trouble."

Tch-29	Subject area/Grade:	English, 10 and 11
	Experience:	25 years
	Gender:	Male
	Discrepancies:	4

"A student's work, participation, and improvement should be the basis of a grade. The student's circumstances, attitude, or discipline problems should not be a factor in an academic grade."

"In all cases, though, a student who makes effort should be given the 'benefit' of the 'doubt' in order to emphasize positive growth."

"The information presented is far from adequate to make a reasonable assessment of a student's work or grade, and therefore, the very premises of this survey approach is ludicrous."

Tch-31	Subject area/Grade:	Chemistry, 10-12
	Experience:	13 years
	Gender:	Female
	Discrepancies:	1

"Every teacher must be able to justify a grade to the student, parent, or administration. It is totally impossible for a teacher to really know every student's situation. Situations that influenced a grade should be dealt with the whole six-weeks. As class sizes remain for many classes 28 or more, meeting the needs of every student is a challenge in which the most hard working, creative, and caring teacher is still a failure due to the demands being so overwhelming."

Tch-33	Subject area/Grade:	English, 11
	Experience:	4 years
	Gender:	Female
	Discrepancies:	1

"I grade according to numbers, not behavior or attitude. I feel my personal rapport with students helps solve a lot of problems."

"I was a bit offended by this survey. I feel that good teachers grade for the work done, not how much, or how little, a student is liked."

Tch-35	Subject area/Grade:	Math, 9-12
	Experience:	24 years
	Gender:	Male
	Discrepancies:	3

"This is not the customary method of assigning grades. It is strictly by numerical average."

Tch-36	Subject area/Grade:	Biology, 10 and 11
	Experience:	18 years
	Gender:	Male
	Discrepancies:	4

"Grading policies eliminate the students' attitudes, etc. we use numbers to calculate grades plus the 'magic pen' comes into use on borderline students."

Tch-37	Subject area/Grade:	Government, 12
	Experience:	25 years
	Gender:	Female
	Discrepancies:	7
	Significant factors:	none

"As teachers get to know and personally work with students, grading patterns shift or change. Personal circumstances of a student often lead me to give them a 'break' - adjust their grades upward a little if there has been a family crisis."

Tch-41	Subject area/Grade:	Social Studies, 9 and 12
	Experience:	21 years
	Gender:	Female
	Discrepancies:	0

"I have always 'awarded grades' on the basis of numerical scores and percentages. To me there does not seem to be room for much subjectivity."

Tch-43	Subject area/Grade:	Algebra II, 10-12
	Experience:	29 years
	Gender:	Female
	Discrepancies:	5
	Significant factors:	none

"Very interesting! I changed my mind after re-reading the instructions and the vignettes, so I have made changes in some of my answers. Hope it is no problem."

Tch-44	Subject area/Grade:	Biology/Chemistry, 10 and 11
	Experience:	4 years
	Gender:	Male
	Discrepancies:	16
	Significant factors:	Behavior and Ability

"This survey of questions does not seem to relate a great deal to my grading policy. I base grades for a student on his performance on tests, quizzes, and other assignments, not how he dresses or problems facing them outside the classroom."

"I circled two grades on several occasions because I was not given the exact numerical value to calculate the grade earned according to the standards I have set up. For example, a 25 is a much lower grade than a 58, even though they are both F's."

Tch-48	Subject area/Grade:	Basic Algebra, 9
	Experience:	9 years
	Gender:	Male
	Discrepancies:	17
	Significant factors:	Effort and Behavior

"I average 1/3 classwork homework, 1/3 test average, 1/3 quiz notebook. Discipline grades are given on the conduct and effort sections of the report cards; however, if a student does the best they can and causes no discipline problem then, if they are on the boarder of failing or passing, I will pass them."

"With the information provided and the way I grade it was difficult to get a full picture of how the student (was) progressing."

Tch-56	Subject area/Grade:	German II
	Experience:	4 years
	Gender:	Female
	Discrepancies:	4

"Inadequate information given concerning homework on some occasions."

"Survey too long for such a short time, one ends up being very approximate".

Tch-57	Subject area/Grade:	Biology, 11
	Experience:	18 years
	Gender:	Female
	Discrepancies:	1

"I always include other mechanisms for grading to reward daily work, project work and laboratory experiences to avoid the 'crisis of grading' of the vignettes. I must substantiate grades so I never depend on only quizzes and tests to really evaluate a student (they would all plummet)."

Tch-59	Subject area/Grade:	Math, 9-12
	Experience:	11 years
	Gender:	Female
	Discrepancies:	4

"If the grades are just judged on tests and quizzes, it's simple. Just average the grades."

Tch-62	Subject area/Grade:	Physics, 12
	Experience:	17 years
	Gender:	Male
	Discrepancies:	11
	Significant factors:	Grade Change

"Calculated the grades by assigning numerical scores to four quiz grades and test. then averaged them with equal weight - per instructions. In my classes quiz and test grades would not have equal weight."

Tch-63	Subject area/Grade:	Foreign Language
	Experience:	5 years
	Gender:	Female
	Discrepancies:	0

"On what planet are teachers allowed to manipulate students' grades based on their perception of the student?? Grade averages are strictly numerically based."

Tch-69	Subject area/Grade:	Government, 12
	Experience:	12 years
	Gender:	Male
	Discrepancies:	2

"The grade is the grade. The time for working with students is before or after the grade is given."

Tch-73	Subject area/Grade:	Biology, 10
	Experience:	2 years
	Gender:	Female
	Discrepancies:	5
	Significant factors:	none

"Do homework grades count at all or is it up to the teacher to decide?"

Tch-77	Subject area/Grade:	Spanish, 10-11
	Experience:	4 years
	Gender:	Female
	Discrepancies:	3

"A grade should not be affected by a teachers opinion of the student. The discipline record should not affect the grade. Grades are grades."

Tch-80	Subject area/Grade:	Chemistry, 10-12
	Experience:	36 years
	Gender:	Male
	Discrepancies:	5
	Significant factors:	Grade Change

"Behavior is not graded in my class."

Tch-82	Subject area/Grade:	Math, 9-11
	Experience:	23 years
	Gender:	Male
	Discrepancies:	4

"Standardized test grades are only indicative of 'potential'. Class performance determines class grade."

Tch-90	Subject area/Grade:	Spanish, 9-12
	Experience:	2 years
	Gender:	Female
	Discrepancies:	7
	Significant factors:	Grade Change

"This was a bit difficult to do without number grades to average. For example, #30 with 3 or 4 grades of B- and one F, the six-week's grade could be a C or D, depending on how low the F."

Tch-91	Subject area/Grade:	U. S. History, 11
	Experience:	23 years
	Gender:	Male
	Discrepancies:	9
	Significant factors:	Effort

"I feel the fact that you like or dislike students can not be a factor in grading. I always try to weight more what a student does best. Test, quiz, homework. I tend in most cases to grade high. I think semester and final grades are most important. Nine weeks grades can be used to encourage or provide a wake-up call."

Tch-94	Subject area/Grade:	English, 12
	Experience:	28 years
	Gender:	Male
	Discrepancies:	11
	Significant factors:	Ability and Grade Change

"Grades made should be given on academic work - not all other."

Tch-100	Subject area/Grade:	Math, 10-12
	Experience:	31 years
	Gender:	Male
	Discrepancies:	10
	Significant factors:	Grade Change

"Percentage of homework, etc. are not given. Degree of difficulty of questions?? More grades than the ones listed are needed so grades way out of balance can be dropped."

Tch-102	Subject area/Grade:	Chemistry, 11
	Experience:	25 years
	Gender:	Male
	Discrepancies:	3

"Homework/classwork were mentioned at various instances, but no weight was given to a grade."

Tch-104	Subject area/Grade:	Algebra II, 10-12
	Experience:	25 years
	Gender:	Female
	Discrepancies:	1

"Quiz grades and test grades would not be given equal weight in my class. I try to grade as objectively as possible; personality problems are considered separately. I try not to let discipline problems influence my grading of a student."

Tch-106	Subject area/Grade:	Spanish III
	Experience:	20 years
	Gender:	Female
	Discrepancies:	2

"(Students in vignettes) 24, 26, 28, 29 would be monitored throughout the year! Their grades would average out."

Tch-107	Subject area/Grade:	U.S. History, 11
	Experience:	24 years
	Gender:	Female
	Discrepancies:	3

"There are some practices outlined in these vignettes that I find disturbing, and a teacher who followed these practices is going

to make his or her life a lot tougher than it has to be. Some changes in classroom management techniques might take care of some of the problems you have included and might also enable students to be more secure and more successful. Of course, there are always some problems that seem intractable.

1. I am very uncomfortable with the number of times that 'talking to other teachers' comes up. A professional should be very cautious about gossiping and chit-chat about students.
2. I am uncomfortable with the almost constant references to standardized test scores. As a student gets older, one has to be very careful about the use of these tests to predict student performance, and in many cases, the scores can become a 'self-fulfilling prophecy.' We should use the scores to help us with specific problems, not as 'predictors' of student success.
3. Discipline problems like talking out of turn or a surly attitude should not be addressed through subjective grading; unless the problems have a direct and measurable impact on a student's work, they should not be 'punished' through grades. There are other ways - and better ways - to handle discipline. Of course, there are times when misbehavior, cutting classes, and the like can directly impact on the student's performance, but I do not think students' grades should be arbitrarily lowered say, five points, because the kids pass notes. There are lots of other ways to handle things like that!
4. There are some very worrisome classroom management problems in these vignettes. In this day and age, a teacher should never leave a class unattended, and leaving a student in charge is almost sheer insanity. I hope the teacher in these vignettes has a really good liability insurance policy.

Note that the teacher 'always' gives quizzes on Mondays and Fridays. I suggest that the teacher take another look at that practice. Of course, there are times when we have to use two days, but students will have higher success rates if the teacher looks at some other days of the week for quizzes and tests.

The teacher seems to rely quite a bit on 'writing kids up' for offenses that could be handled in more effective ways. The teacher who relies constantly on 'writing up' to handle discipline will soon find (a) an administration that does not take his/her problems seriously or (b) students who do not take 'writing up' seriously. Save the 'writing up' for the big stuff. for example, in one case a student cursed at the teacher - that should be 'written up,' just as cutting class should be written up.

I must object most strenuously to one teacher practice that appears and reappears throughout your vignettes. I was shocked to read the teacher did not grade assignments. Many problems will inevitably stem from this practice. In the first place, the teacher will very quickly find that students don't do the assignments, or will simply copy them from some poor sole who does. The students who do the assignments will become demoralized. If the assignment is worth giving the teacher has AN OBLIGATION to read it and give the student some sort of credit for the effort he expends on it. If the assignment is not worth the teachers time, why should the student be expected to do it? Years ago when I was in teacher training my professor at William and Mary told me, 'If you assign the work, you must do your part.' For him, it was a matter of integrity, and I have never gone wrong by following his advice.

On a related point, I think it is unfair to base a student's grade solely on quizzes and tests and to take no account of work completed in class, homework, and daily grades. There are lots of good ways to give credit for work done in class and to give extra credit for class participation. These types of 'grades' can really help kids, especially kids in trouble, kids with family problems, kids with jobs, kids with no quiet time to study.... These grades can serve as 'insurance policies' and they give lots of chances for constructive feedback and support. Kids who know their efforts will pay off in a positive way are much more likely to try on difficult assignments, to complete assignments, and to use the time in class constructively. When a teacher has no way of marking that sort of work, the students have little incentive to prepare for class or use class time wisely. As a result, classroom management problems can rapidly increase.

If you are interested, I can show you a couple of simple tricks of the trade (borrowed from others, adapted from others) that have really helped me over the years.

Tch-109	Subject area/Grade:	Chemistry, 10-12
	Experience:	8 years
	Gender:	Female
	Discrepancies:	0

"Hard to do. Sometimes background does influence grades. without numbers and seeing the quiz grades (not just their average) it was hard to feel comfortable about the grades I decided on."

Tch-110	Subject area/Grade:	Math, 10
	Experience:	25 years
	Gender:	Male
	Discrepancies:	0

"With no set criteria for homework it is hard to make a judgement. Also, one should not let outside forces influence grading procedure. My assumption is that most school systems do not allow for poor home situations, jobs after school, etc. to affect grading. One, pretty much gets what he earns. Sometimes there are unfortunate situations, but that should not intervene. The same goes for class clowns, class pains, and class nice people. Not to sound harsh or robotic, but you should always be objective in grading. It is the only fair way to be. If not, you will ultimately cheat someone, whether by accident or on purpose."

Tch-111	Subject area/Grade:	U.S. History, 11
	Experience:	3 years
	Gender:	Male
	Discrepancies:	8
	Significant factors:	Effort and Behavior

"I consider homework and classwork important indicators of effort put forth by students. A slower student who works hard should be given credit for his effort, while a bright student should not be allowed to 'slide by' without putting in effort. I also feel that bright students need to be held accountable for work so that

they learn good work habits for when they reach the point where they have to study."

Tch-113	Subject area/Grade:	Chemistry, 11
	Experience:	7 years
	Gender:	Female
	Discrepancies:	1

"I believe students should receive the grades they earn, but I always discuss the progress with them. We talk about problems at home and how they cannot allow them to affect their school work. I offer extra help and keep after them to complete and turn in work. If they understand why they got the grade they got and that you are concerned about them and want to help them improve a low grade does not have to be detrimental. It can actually be beneficial to their future performance more so than an artificially inflated grade. I do not feel that you can accurately grade a student's performance on tests and quizzes only. If grades were taken on homework and class participation it would give hard workers and poor test takers a better chance and make those who do well on tests, but have bad attitudes, improve their attitudes if they want to maintain their good grades."

Tch-120	Subject area/Grade:	Physics, 11-12
	Experience:	23 years
	Gender:	Female
	Discrepancies:	5
	Significant factors:	none

"To me, behavior should not really influence grades very much. Of course you are concerned about bad behavior, but you work on it in another way."

Tch-123	Subject area/Grade:	English, 12
	Experience:	29 years
	Gender:	Female
	Discrepancies:	7
	Significant factors:	Effort and Behavior

"It appears there are more C's than any other grades."

Tch-125	Subject area/Grade:	Spanish, 10-12
	Experience:	18 years
	Gender:	Female
	Discrepancies:	5
	Significant factors:	none

"These vignettes are unrealistic in that grades are an evaluation of a student's progress at a certain point in time. It is not related to behaviors that are considered disruptive or a problem. A grade is not an isolated factor. a grade is to show a student where he is and what he needs to do to master the skills necessary for the most positive and complete life style possible."

"I do not teach a subject, but a skill so these vignettes do not relate at all to my situation."

"This exercise seems to be an academic excuse for a graduate student to justify his existence. It does not relate to the real world."

Tch-126	Subject area/Grade:	Social Studies, 9-12
	Experience:	21 years
	Gender:	Male
	Discrepancies:	10
	Significant factors:	Behavior

"Usually in my grading, I assign simple reports as mandatory work for test or quiz credit. I also use class behavior as a major consideration in final grading."

Tch-127	Subject area/Grade:	U.S. History, 11
	Experience:	12 years
	Gender:	Male
	Discrepancies:	2

"Grades are grades."

Tch-131	Subject area/Grade:	Biology, 10-12
	Experience:	8 years
	Gender:	Female
	Discrepancies:	2

"The general affect of a student (attitude, prior discipline, physical appearance) cannot come into play when grading. I feel strongly that I must be able to back up grades I have given with numbers."

Tch-133	Subject area/Grade:	English, 11
	Experience:	4 years
	Gender:	Female
	Discrepancies:	13
	Significant factors:	Behavior

"This exercise really made me take a close look at how I grade students. There are many things other than actual academic ability that must be taken into consideration when assigning students a particular grade. Effort of course must be considered, as well as, the situations most students live i at home. As teachers we should not be out to totally discourage students. We want them to feel good about themselves but we do not want to give them a fake sense of where they stand either. I liked doing the exercise but it was tough! I felt like I should be giving some +'s and -'s! I guess the reason for that was because of certain motivational attitudes put forth."

Tch-134	Subject area/Grade:	English, 11
	Experience:	29 years
	Gender:	Female
	Discrepancies:	3

"Grades are earned, not assigned."

Tch-135	Subject area/Grade:	Government, 12
	Experience:	24 years
	Gender:	Male
	Discrepancies:	1

"Should have been given the option of a + or - on the grades assigned. Not enough grades were taken to arrive at a valid

assessment. Other types of grades other than tests and quizzes should be an option."

Tch-139	Subject area/Grade:	Spanish, 10-12
	Experience:	7 years
	Gender:	Male
	Discrepancies:	1

"With no daily grades or homework grades to take into account and no numerical grades to go by, this whole thing is hit and miss guesswork."

Tch-140	Subject area/Grade:	English, 10 and 12
	Experience:	25 years
	Gender:	Female
	Discrepancies:	11
	Significant factors:	Grade Change

"Grades have to be earned and not assigned. Six week tests should be worth more than a daily quiz."

Tch-142	Subject area/Grade:	Earth Science, 9
	Experience:	10 years
	Gender:	Female
	Discrepancies:	9
	Significant factors:	Effort

"I use +'s and -'s tacked onto grades to differentiate borderline situations. Homework and classwork are numerically accounted for in my classes. They literally factor into the student's average. You seem to be implying that grades are a judgement call. I do not feel comfortable with this. In my classes assignments are graded or I do not count them in student's averages. I very rarely factor all of the other parts of the scenario into a grade. It might affect how much help or encouragement I offer to better the student's performance. I assumed the vignette students finished (their) homework and classwork."

Tch-145	Subject area/Grade:	Biology I, 9-10
	Experience:	16 years
	Gender:	Male
	Discrepancies:	0

"Quizzes count 50% test counts 50% - the rest is basically unnecessary information."

Tch-149	Subject area/Grade:	English, 12
	Experience:	20 years
	Gender:	Female
	Discrepancies:	8
	Significant factors:	Behavior

"In our county, .5 always rounds up as I understand it. In a real situation, there would be other considerations: homework grades, participation, etc. which would be assigned more arbitrarily than number grades. In those areas I would give more consideration to home life, behavior, attitude, and ability tests."

Tch-152	Subject area/Grade:	Earth Science, 9
	Experience:	35 years
	Gender:	Male
	Discrepancies:	1

"With only one test grade we can be catching the student on a bad day, however I did not make any such assumption. In my own classes I do not box myself in with only two sets of grades. Daily homework averages and projects, labs, etc. help to get a better look at their ability."

Tch-156	Subject area/Grade:	English, 9
	Experience:	20 years
	Gender:	Female
	Discrepancies:	5
	Significant factors:	Grade Change

"(Vignettes) Number 3, I would review other aspects and perhaps give a B, a tough one! Number 6, either C or D...hard again! Perhaps a D would spur her next time. Number 7, also tough. Number 9, talk with her and hope to 'inspire'. Number 10,? Number 23, hard call. Number 28, C+ ? Tough one."

Tch-162	Subject area/Grade:	English, 10
	Experience:	17 years
	Gender:	Female
	Discrepancies:	6
	Significant factors:	none

"I try to grade fairly in all my classes no matter if a student is a problem or not. Students will learn later in life that they need to study, work hard, and pay attention."

Tch-171	Subject area/Grade:	English, 9
	Experience:	21 years
	Gender:	Male
	Discrepancies:	4

"Homework is an ingredient missing in most. Why?"

"Effort counts on my FINAL analysis!"

"Personalities do not effect grades I give."

"I assign little homework besides reading."

Tch-172	Subject area/Grade:	English, 11
	Experience:	1 year
	Gender:	Female
	Discrepancies:	8
	Significant factors:	Behavior

"I also have a problem with the emphasis placed on the standardized tests - I do not think they should be considered in so important terms. Most of the time teachers do not know the student's scores and those tests are biased and un-cultural."

Tch-175	Subject area/Grade:	Algebra I, 9
	Experience:	4 years
	Gender:	Female
	Discrepancies:	5
	Significant factors:	none

"Discipline problems and peer choice should have no affect on grades. Cheating receives a zero which affects that grade. Who cares if you like a child or not, WE HAVE A JOB TO DO!!!!"

Tch-185	Subject area/Grade:	Math, 9
	Experience:	9 years
	Gender:	Female
	Discrepancies:	11
	Significant factors:	Grade Change

"Daily work to me is more important than either quizzes or tests. A poor grade on a test or quiz may indicate a difficult day. however, you did not give any idea about daily grades."

Tch-192	Subject area/Grade:	Earth Science, 9-10
	Experience:	10 years
	Gender:	Male
	Discrepancies:	4

"You should have said low F or high F, as you did with the +'s and -'s. I quit reading the stories and focused in on the grades only. I do give the benefit of the doubt when a student tries when they are between two grades. If they do not try, then they get the grade of the number earned."

VITA

FRANKLIN A. BRUCE, JR.

**2509 Gloucester Drive
Blacksburg, VA 24060
Home: (703) 953-3141**

EDUCATION:

- January 1995 Doctor of Philosophy, Educational Research and Evaluation, Virginia Polytechnic Institute and State University
- February 1988 Master of Science, Special Needs for Vocational Education, Virginia Polytechnic Institute and State University
- June 1973 Bachelor of Science, Industrial Arts Education, Virginia Polytechnic Institute and State University

PROFESSIONAL EXPERIENCE:

- May 1994 - Research Assistant, National Center for Research
December 1994 in Vocational Education, Berkeley, California

Responsibilities: Planned and conducted interviews with vocational leaders nationwide and qualitatively analyzed data.

- August 1991 - Testing Coordinator for Montgomery County Public
July 1993 Schools, Christiansburg, VA

Responsibilities: Coordinated county and state assessment programs to include state's Literacy Testing Program (LTP). Interpreted state, district, and school test score reports.

Assessment Specialist for Programs for the Gifted

Responsibilities: Coordinated and assessed nominees for Programs for the Gifted. Tests administered and interpreted included: CogAT (Cognitive Abilities Test), OLSAT (Otis-Lennon School Ability Test), TCS (Tests of Cognitive Skills), GIFT

(Group Inventory for Finding Creative Talent), ITBS (Iowa Test of Basic Skills) and a interest inventory - SII (Student Interest Inventory). Developed a data base of gifted nominees' assessment information using dBase IV.

Summer 1988 - Graduate Assistant for Research Laboratory,
Spring 1991 Educational Research and Evaluation, Virginia Polytechnic Institute and State University

Responsibilities: Conducted seminars and assisted students in class, thesis, and dissertation analyses. Used and taught PC software: Number Cruncher Statistical Systems (NCSS), Stat Pac Gold, Word Perfect; also, assisted with SPSSX and SAS.

1988 - 1989 Graduate Research Assistant Roanoke City Schools Evaluation Project, Educational Research and Evaluation, Virginia Polytechnic Institute and State University

Responsibilities: Assisted in developing surveys, sampling plans, and coordinating data collection using telephone and mailings to seniors, graduates, teachers, parents, classified personnel, business/industries, and administrators. Also, wrote and edited summary reports.

1987 Graduate Assistant Gender Equity Project, Vocational and Technical Education Division, Virginia Polytechnic Institute and State University

Responsibilities: Assisted in writing newsletters and coordinating seminars.

TEACHING EXPERIENCE:

August 1984 - Ungraded Program Teacher Prince Edward County
June 1986 Elementary Schools, Farmville, VA

Responsibilities: Developed and taught remedial math, industrial arts, and physical education to economically and culturally disadvantaged early adolescent students.

August 1982 - Coordinator of Industrial Cooperative Training
June 1983 Charlotte County Schools, Charlotte Court House, VA

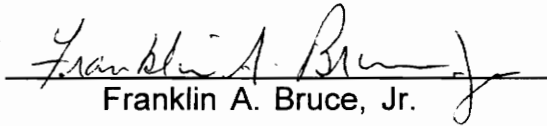
Responsibilities: Supervised and coordinated work release job experiences. Taught career educational classes.

January 1982 - Industrial Arts Teacher Prince Edward Academy
June 1982 Farmville, VA

Responsibilities: Instructed high school students in machine woodworking, metal work, and drafting.

August 1974 - Industrial Arts Teacher Montgomery County
June 1978 Schools, Christiansburg, VA

Responsibilities: Developed an Industrial Arts Program and taught regular and special education students grades 6 - 9.


Franklin A. Bruce, Jr.