

**BIAS IN THE ASSESSMENT OF WRITING**

**KEITH HARRY FISHER**

B. Ed., University of Calgary, 1974

A One-Credit Project  
Submitted to the Faculty of Education  
of The University of Lethbridge  
in Partial Fulfillment of the  
Requirements for the Degree

**MASTER OF EDUCATION**

LETHBRIDGE, ALBERTA

September, 1990

## Acknowledgements

I would like to express my appreciation to Dr. Anne Murtagh for the very valuable feedback she provided for each draft of this project, and for her encouragement and her enthusiasm for the project.

I also appreciate very much the assistance and helpful suggestions provided by Dr. Ritchie Whitehead.

I am grateful to the students in my English 30 "Class of 1990" for participating so enthusiastically in the study.

My wife, Cheryl-Ann, and my children, Barry and Carla, have been extremely patient and understanding while I have been occupied with my studies; I love and appreciate them very much.

## TABLE OF CONTENTS

INTRODUCTION .....	1
The Evaluation of Writing .....	1
The Difficulties of Evaluation .....	4
Statement of the Problem .....	9
RELATED LITERATURE .....	10
Characteristics of Markers .....	10
Characteristics of Writing .....	13
Characteristics of Writers .....	15
DESCRIPTION .....	16
Introduction .....	16
The Sample .....	16
The Assignments .....	18
Procedure .....	20
ANALYSIS OF PRE-SURVEY .....	25
ANALYSIS OF SCORES .....	29
ANALYSIS OF POST-SURVEY .....	34
FINDINGS AND CONCLUSIONS .....	41
Analytical Comments .....	41
Interpretive Comments .....	43
Summary .....	47
REFERENCES .....	
APPENDIX A Reading List .....	
APPENDIX B Pre-survey .....	
APPENDIX C Post-survey .....	
APPENDIX D Letter .....	
APPENDIX E Scoring Guide .....	

## LIST OF TABLES

1.	Responses to Pre-survey	27
2.	Scores in % For Each Essay	30
3.	Means of Essays Without Names and With Names	32
4.	T Test of Significance of Difference Among Scores of the Two Marking Strategies (With Names and Without Names)	33
5.	Scores For Essays With Names and Without Names For the Top Seven Students Ranked According to Scores For Essays With Names	34
6.	Scores For Essays With Names and Without Names For the Bottom Seven Students Ranked According to Scores For Essays With Names	35
7.	Responses to Post-survey	36

## BIAS IN THE ASSESSMENT OF WRITING

### Introduction

#### The Evaluation of Writing

Teachers who evaluate student writing have before them a difficult task. The teacher must somehow assign a number to a piece of writing which represents, in a personal sense, a student's feelings and experience, and which represents in a cognitive sense, the student's ideas, philosophy, research and thoughts. Students' essays are, then, much more than a mere collection of words and paragraphs. Teachers must be sensitive to both the personal element which is inherent in the work they evaluate and to the cognitive aspect of the work. Cooper and Odell (1977) note that since "writing is an expressive human activity, we believe the best response to it is a receptive, sympathetic human response." At the same time, the process of evaluation must be a fair one, and therefore teachers must focus their assessment on the work itself, not on the student. Because of the personal nature of writing, students leave themselves in a vulnerable position when they submit written work for marking. They may feel that an evaluation of their work is also an evaluation of the writers themselves. To some extent, this statement is true. After all, what people write is a product of their

experience in life. How they write is a product of how they feel and think. Secondly, because of the cognitive aspect of writing, students may feel that an evaluation of their writing is really an evaluation of their intelligence and their creativity. This statement, too, contains some truth. What teachers want to see in student writing is a story which is interesting, an essay which is personal, a poem which is creative. All of these things require thought and reflection. Without cognitive skills, writing may be a frustrating task.

At some point, the teacher must evaluate students' writing. As previously noted, a close relationship exists between the writer and the writing, and one of the major difficulties in evaluation is to separate the two. The attempt at separation is largely an attempt to remove bias (either positive or negative) from the evaluation.

The essays students submit to their teachers are usually a culmination of many kinds of activities. Cooper and Odell (1977) note some of these activities in their description of the process of writing:

Composing involves exploring and mulling over a subject; planning the particular piece (with or without notes or outline); getting started; making discoveries about feelings, values, or ideas, even while in the process of writing a draft; making continuous decisions about diction, syntax, and rhetoric in relation to the intended meaning and to the meaning taking shape; reviewing what has accumulated, and anticipating and rehearsing what comes next; tinkering and reformulating; stopping;

contemplating the finished piece and perhaps, finally, revising. (p. xi).

All of these considerations have implications for the teacher who must evaluate writing. When a teacher evaluates a composition favorably, the students, naturally, feel good about themselves. However, if the evaluation is unfavorable, students often feel that they have been personally criticized or violated. It may be difficult for them to separate the writing from the writer. A second implication relates to the teacher. If writing is linked so closely to the writer, what, exactly, should be evaluated, and how should this evaluation be achieved?

Teachers of mathematics rarely face this problem, because the answers to the problems will be clearly right or wrong, and the "exposing" of one's thoughts and feelings is usually not involved. To return to the question, teachers who must evaluate writing usually do not focus on a perceived arrival at the "correct" answer. Louis Rosenblatt holds that we should reject the "preoccupation with some illusory unspecifiable absolute or 'correct' reading..." (Rosenblatt, 1978). For example, in an English 30 class, students do not have to interpret Hamlet's soliloquies in exactly the same way as the teacher. To this writer, the objective is to have students engage in the processes of thinking about what Hamlet says and feels,

and then to come to an intelligent, logical and defensible understanding of the character. However, these qualities are, to a large extent, intangible. As a result, the process of evaluation becomes further complicated. The assessor must judge the overall quality of the work by considering the intelligence of the ideas, the logic of the arguments and the evidence which supports the ideas. Also, since the goal of writing is to communicate, the teacher must evaluate the student's ability to use language effectively. To have good ideas is half the battle; the other half is to be able to clearly communicate the ideas. This duality is of necessity important for evaluators: they must appraise both the quality of the content of the work, and the linguistic skills of the writer.

#### The Difficulties of Evaluation

The aforementioned conditions create certain difficulties in assigning marks to a piece of writing. One such difficulty is that the assessment of writing is a subjective activity. Romano (1987) makes this point quite plain:

Evaluation of writing is a necessarily subjective act. Objectivity is impossible. Participate in one group grading session and you'll realize that. When many teachers evaluate the same paper, their judgements of its merit are diverse and astounding. So I am left with my subjectivity. (p.16).

The research literature suggests that many factors may

increase the level of subjectivity, to the point where the evaluation of writing is not valid and not reliable. The factors may generally be divided into three groups: those which relate to the characteristics of the compositions, those which relate to the characteristics of the evaluators and those which relate to the characteristics of the writers themselves.

The characteristics of the compositions may influence the evaluator positively or negatively. One such feature is appearance. For example, a teacher usually will react positively to a composition which is double-spaced and which contains neat handwriting. The effect may be pronounced if the composition appears near the end of a marking session, when the teacher is tired, or if the composition follows several essays which are practically illegible.

Another characteristic of the writing which may influence the marker is the presence or absence of mechanical errors. This category includes errors in spelling, errors in subject-verb agreement, errors in pronoun-antecedent agreement, errors in punctuation and errors in capitalization. Teachers may assign a relatively low mark to a piece of writing which contains many mechanical errors, even if it also contains thoughtful and intelligent ideas.

The style of writing and subject nature of a composi-

tion also may influence evaluation. Student compositions which are humorous, ironic, satirical, creative or unusual in any way may be received very favorably by teachers who value creativity and experimentation. These teachers will also reward compositions which deal with controversial topics.

Characteristics of evaluators are also a consideration in the evaluation process. A teacher who is rigid and intolerant might punish a student who writes creatively or who writes about controversial topics. This kind of approach is unfortunate, because, in this writer's mind, it limits creativity and discourages the student from independent thinking. Other teachers might reward students who agree with the teacher about issues discussed in class. The students soon discover that to succeed, they must imitate the teacher's ideas. Again, the quality of the learning experience suffers, and the evaluation process becomes flawed.

Teachers may also react to the way various students are "labelled." For example, a student who usually receives "straight A's" may receive high marks even when the work is average. A student who usually barely passes may not receive a high mark even when the work is proficient. Students who are cognizant of this characteristic of their teachers often try to take advantage of it by submitting

superior work during the first month of the semester, and then hoping the academic "label" will carry their marks for the duration of the term. The situation may work in reverse for the weaker students who initially receive low marks, and then feel that there is no hope for improvement.

Teachers may be victims of the "halo effect." This occurs when the marker responds to and focuses on one aspect of a composition and allows that reaction to sway the mark either upward or downward. For example, teachers may respond to a brilliant introduction and give a high mark, even if the writer fails to adequately develop a thesis.

Teachers who mark holistically also run the risk of producing an invalid evaluation. Holistic marking involves reading a piece of writing and basing a mark on the composition as a whole. The rationale for the mark is not evident.

Several unacknowledged biases may be at work when teachers evaluate their students. The teacher may be influenced by factors such as the quality of the relationship between the teacher and student. Some of the indicators or contributors to this relationship include the quality of talk between the teacher and student. For example, one of the indicators is whether or not the teacher says "Good morning" to all the students, some of the stu-

dents or none of the students, and whether the students respond to or even initiate the greeting. Another consideration may be how close or far away the student sits from the teacher's desk. Students who have interests or ideas which are similar to the teacher's may also benefit when their work is evaluated. The teacher's relationship with parents probably is another important consideration. If the teacher knows that parents of a certain student have high academic expectations, the teacher may feel under some pressure to give marks which meet those expectations.

Other characteristics in evaluators include intra-rater reliability and inter-rater reliability. The former refers to agreement of a rater with himself at different points in time. For example, if a teacher marks a set of essays over a period of a few days, and becomes more demanding as he marks, his reliability suffers. Inter-rater reliability refers to agreement among different raters. Different teachers may assign significantly different marks to the same essay. Some of the contributors to low inter-rater reliability include the teachers' experience, their knowledge of writing and their expectations for student writers at various points of the school year.

Students who socialize during class may have their work marked more closely and rigorously than students who

appear to the teacher to be attending to the assigned work. The students' behaviour influences the evaluation.

The characteristics of the students themselves are the third component which may influence evaluation. Some students may be luckier than others when they are evaluated. Their names may draw a positive or negative reaction from the teacher. Also, their appearance may influence their evaluation. Students who are attractive and neatly dressed might have an advantage over those who are unattractive or who dress in ways which the teacher dislikes.

Even at the best of times, evaluation is not an easy task. White (1988) notes that "we can never determine a student's true score on a test." This comment underscores the difficulty of the valid and reliable assessment of writing. Archer and McCarthy (1988) acknowledge that bias does exist in writing assessment, and recommend that further research be aimed at "assessing the extent of bias in real, as opposed to simulated, assessments." This leads to the nature of the present study, which uses actual student essays.

#### Statement of the Problem

The purpose for the study was to check for any bias which may be present in my evaluation of student writing when identities are known. The study is designed to investigate the question: Do students' scores on essays change

significantly when their identities are removed from the essays during the evaluation process?

### Related Literature

#### Characteristics of Markers

As noted earlier, characteristics of the markers may affect assessment. For example, a discrepancy may exist between what teachers say they evaluate, and their actual evaluation. One of the most frequently cited conditions is the "halo effect." Here, teachers respond to one feature of a composition, but ignore all the other features. Gronland (1981) notes that the halo effect may be caused by "a certain general impression made by the first few answers to the questions." It may also be caused by the teacher's impressions of the student. To control the halo effect, Gronland recommends that teachers assess all answers to each question at one time and to evaluate without looking at the students' names.

In a study of Diploma Examinations in Alberta, Nyberg (1988) suggests some markers are influenced by a halo effect. When markers liked or disliked one aspect of a composition, they tended to reward or penalize the writer on all the scoring variables.

In a study about the effect of students' personal qualities on their evaluation, Wade (1978) found that

students' first names influenced teachers when they assessed assignments. For example, if a teacher has several students by the name of "Linda" who are superior students, there may be a carry-over. The teacher's expectation for the next "Linda" will be that she too will be a very capable student. Wade also found that teachers were influenced positively by neat handwriting and by the attractive physical appearance of particular students.

Diederich (1974) also notes the presence of bias in evaluation. He says that some teachers react strongly to certain errors. Further, teachers are most biased when they mark the compositions of their own students; they are likely to change too little when the essay of a given student surprises or disappoints them. Even more alarming, Diederich found that when papers were stamped in an alternate, random manner, with a "regular" or "honors label," the papers stamped "honors" averaged almost one grade-point higher than copies of the same papers stamped "regular."

Related to these findings is a study by Graham and Dwyer (1987). They also take up the issue of bias which results from knowledge of academic aptitude. Evaluators were informed of the educational status (learning-disabled and normal) of the subjects and were then asked to mark essays. Results suggested that markers were influenced by

the labels, but that training and practice helped reduce the effects of knowing of the students' abilities.

McCaig (1982) points out the necessity of evaluator reliability. He found that not all teachers can become good judges of writing, even with adequate training. Some were "simply found not to be knowledgeable analysts of written language." Others could not free themselves from "deeply ingrained attitudes and rules even though they often professed different beliefs." Again, the discrepancy between what teachers claim to consider important in evaluation and what they actually examine becomes apparent.

Jolly (1985) also focuses on the role of the evaluator. She observes that "the majority of writing instructors act out of ignorance. They simply do not know how to evaluate a piece of writing." She recommends looking at content, organization, grammar and mechanics, and further suggests that the paper be looked at as an entire unit. Focusing on these components of writing does assist in reducing bias.

Another form of bias relates to ethnic and cultural differences between the students and the evaluators. Chodzinski (1988) refers to a 1985 Ontario Ministry of Education Race Relations Committee study which acknowledged that minorities faced a difficult task in

participating fully in Canadian society. One of the results of this study was a policy which states "Racism in any form will not be tolerated in Ontario schools." Unfortunately, the problem of racial and ethnic bias in education is real. Chodzinski draws attention to the reality that minority children are often streamed to non academic programs. He said that part of the problem is that teachers have no training, no support and no inservice to help them deal with students of different cultures. This topic will become of even more concern to educators in Canada through the 1990's as more people immigrate to this country.

#### Characteristics of Writing

As mentioned in the introduction of this report, the second contributor to subjectivity relates to the characteristics of the writing itself. The appearance of the essay can influence how it is evaluated. Sloan (1978) conducted a study to investigate the effects of handwriting on the grades given to essays. He had some essays randomly selected and rewritten by Palmer handwriting experts. He concluded that papers written in expert handwriting were evaluated by the teachers as "significantly better papers."

Peacock (1988) compared handwriting to word processed

print. Trained teachers marked essays which were handwritten. Other trained teachers marked the same essays, but in word processed form. Higher grades were assigned to the word processed essays than to the handwritten ones. This suggests that the appearance of the assignments does influence the evaluation.

The effect of quality of handwriting, spelling accuracy and use of a scoring key on essay scores was investigated by Chase (1968). He concluded that the quality of handwriting has a "significant influence on scores." Students who wrote neatly reaped dividends. Further, Chase found that spelling is not significantly related to scores. Many people who are interested in writing assessment often fear that poor spelling contributes to poor marks, but Chase disproves this assumption. Chase also noted that markers who used a scoring guide tended to give a paper higher marks than readers who did not have a guide.

Marshall and Powers (1969) arrived at similar conclusions about the effects of handwriting on scores.

The actual content of written works also influences the evaluators. Freedman (1979) manipulated four characteristics in essays: content, organization, sentence structure and mechanics. She then rewrote essays of moderate quality to be either stronger or weaker in the above

four areas. She found that the most important influences on scores were content and then organization.

In a similar study, Freedman and Calfee (1983) found results which paralleled those of Freedman's earlier studies. They concluded that idea development influenced markers the most, and mechanics influenced markers the least. Similar results were reached by Breland and Jones (1984).

Ed White (1988) notes that one of the problems with scoring essays is that it is impossible to determine a student's true score on a test. He recommends that essay test administrators should "reduce the sources of variability in test contexts, should keep the scoring criteria constant, should pretest and control test prompts, should control essay reading and scoring procedures, and should always try to use multiple measures to assess students' skills." These strategies would indeed increase validity, but they would obviously be impossible for teachers to employ on a regular basis. It is quite easy to keep scoring criteria constant by using a scoring guide, but teachers may wish to use different scoring guides for different types of writing. Using multiple measures for assessing students' skills is an appropriate strategy, but may be too time consuming to be realistic.

#### Characteristics of Writers

Archer and McCarthy (1988) discuss bias extensively,

and note that bias "can result from knowledge of a person's sex, physical attractiveness and past assessment performance." Their research, and the studies they have examined, point to the fact that students are not evaluated fairly; some students will be marked easier because of their socioeconomic background, because of their physical appearance, or because of their past success at school.

### Description of Study

#### Introduction

As noted earlier, the purpose of the study was to investigate the question: Do students' scores on essays change significantly when their identities are removed from the essays during the evaluation process?

Students wrote six essays during the study. Three of these had names attached, and three had numbers attached. The means of essays without names were compared to the means of essays with names.

#### The Sample

The sample used in this study consisted of the writing assignments of a Grade 12 English 30 class. English 30 is a matriculation course and is required for university entrance. English 10 and 20 are prerequisite courses. These courses are bipartite in nature: they combine a study of

literature with a practice in composition. Alberta Education specifies in the 1982 Senior High School Curriculum Guide for Language Arts that the content of English 30 include at least six short stories, eight nonfiction articles, one modern play, one Shakespearean play and a selection of poetry. The composition assignments in English 30 are of two types: personal response and critical analysis. A typical personal response topic requires students to respond to a piece of literature by choosing a quotation from the selection and explaining why this quotation is meaningful to the student. Students may write in any genre including diary, exposition, description, narration, poetry and letter. The critical analysis requires students to analyze literary selections mainly in terms of themes and ideas. Other literary elements such as characterization, plot, setting, symbol, irony and metaphor usually become part of the supporting discussion.

The class itself consisted of twenty-five students, including sixteen females and nine males. Twelve of these students were not new to me - they had previously attended my class in Grade 10 or 11. At the beginning of the semester, two of these students were initially assigned to another teacher, but they requested a transfer to my class. Their rationale for transferring was that I knew them and that they were aware of my expectations. They

wanted to be successful, and thought that this move would enhance their achievement. They realized that there is a time constraint inherent in the semester system, and to them it was advantageous to begin the course with the teacher-student relationship already developed.

The students' prerequisite marks indicated that of the twenty-five students, five were very strong academically; their marks for English 20 were higher than 80%. Another nine students were also strong; their prerequisite marks were in the 65-79% range. The remaining students' marks, then, were located in the 50-64% range.

### The Assignments

The six essays which were assigned were of the critical analysis genre. For the first essay, students had to read a novel from a reading list (Appendix A). Titles for this list came from Alberta Education's recommended list and from my own reading background. Students could also read a novel not on this list as long as I approved it. The assignment for the first essay was to choose three passages and discuss the significance of each one and to make connections among the passages in terms of development of character and theme. One passage had to come from each third of the novel. This assignment forced students to "make connections" among various parts of the novel,

and hopefully to come to a synthesis of the meaning of the work.

The second assignment was a comparison/contrast of two short stories. Students read "Young Goodman Brown" by Nathaniel Hawthorne and "The Lottery" by Shirley Jackson. Following class discussions and small-group sessions where students made notes and engaged in further discussion, the topic was assigned. This topic asked students to compare and contrast the ideas in each story, and to look at techniques used by each author to develop the ideas.

The third essay related to Henrik Ibsen's play A Doll's House. The assignment was to identify some of the general ideas developed in the play. For example, one of the ideas could be that an artificial relationship between a husband and wife will not endure. Another idea might be that a sense of personal identity and personal worth is very important.

Students next looked at a modern novel. This was Judith Guest's Ordinary People. The assignment here was to write about some of the themes which Guest develops. These themes included coping with a death in the family, the grieving process, coping with loss, coming to terms with feelings of guilt, acceptance of self, sex and sexual relationships and reconciliation. This novel is controversial because it deals with teenage suicide and uses some

profane language. It is recommended by Alberta Education and students find it interesting and meaningful.

Following the fourth essay, students viewed the film version of Ordinary People. They worked through a study guide I have written, and then were given a choice of three topics. One related to censorship, and asked students to consider the controversial nature of the content and to decide if the novel was worthwhile for high school students. The second choice required students to examine the film director's interpretation of the novel, and to look specifically at certain characters and scenes. The third choice involved a comparison of the novel and the movie, and asked students to choose which medium they considered more effective.

The final essay followed a study of three sonnets: "Sonnet 17" by George Meredith, "Sonnet 116" by Shakespeare and "Sonnet 18," also by Shakespeare. Students had to compare and contrast any two of the three sonnets. Like most of the previous essays, this activity meant students had to identify significant ideas in the literature, and then look at the authors' use of literary techniques to develop the ideas.

### Procedure

Data from a pre-survey (Appendix B), from six essays

and from a post-survey (Appendix C) was collected for analysis. At the beginning of the semester, the nature of the study - to compare the scores for essays with names to the scores for essays without names - was explained to students and a letter to parents (Appendix D) was sent home with them asking for permission for each child to participate in the study. All students were able to participate. Before the students did any writing, they completed the pre-survey. This pre-survey was placed in a sealed envelope which was not opened until the study was completed. Students were informed of this procedure, and reminded that their marks would not be jeopardized because of anything they wrote on the survey forms.

Students used computers in the school to write all the essays. This strategy allowed all essays to look the same. Students did not attach their names to the first, third and fifth compositions. Instead, they used a five-digit number of their own choosing. They did not use the same number for all three unidentified essays. As a result, there was no way for these three essays to be identified during the evaluation process. To return the marked essays to the students, the identification numbers were read aloud in front of the class, and students came forward to claim their essays. Students examined their essays, put their names on them and returned them to me. Scores were

then recorded in the mark book. Students could check to be sure the recorded marks were accurate. The essays were then put into individual writing folders, and kept in the classroom for the duration of the semester. Students could access these folders at any time.

As described earlier, all essays required students to critically analyze the literature they studied in class. This procedure increased the validity of the results, since the same type of writing was evaluated for each essay.

The students wrote the essays over a period of fourteen weeks. They had three class periods of eighty minutes each to work on each essay. In addition, the computer lab was available at noon and after school for students who required additional time to complete their assignments. While working on their essays, students were free to consult with their peers and with me. Various writing skills were taught and reviewed throughout the semester, particularly in the days before an essay was assigned, and when essays were returned. Some of these skills included various methods of writing an introduction, writing a conclusion, using a variety of sentences, using various types of sentence for particular effects, developing vocabulary, developing tone and voice, and supporting ideas and interpreting the supporting evidence.

Students used "Word Perfect 4.2." Most of them had used this program in previous computer literacy classes. The other students experienced very little difficulty learning to use the program, and, in fact, received assistance from the "computer literate" students whenever they needed it.

All essays were evaluated with the same instrument: the "English 20 and 30 Critical Response Scoring Guide" (Appendix E). This guide is modelled after the Alberta Education Diploma Examination Scoring Guide. The major difference between the two is that Alberta Education has a "Total Impression" category which I do not use. The guide used in the study assesses four areas of writing: "thought and detail," "organization," "matters of choice" and "matters of convention." Each area has a four-point scale: excellent, proficient, satisfactory and deficient. Excellent corresponds to an "A," proficient to a "B," satisfactory to a "C," and deficient to a "F."

"Thought and detail" relates to the quality of ideas. The descriptor for the "Excellent" category notes "insightful ideas are supported by carefully chosen details. Literary interpretations are perceptive and defensible." The "Deficient" category says "ordinary ideas are weakly supported...and literary interpretations are incomplete or superficial." To receive high marks for thought and de-

tail, the student must carefully and thoughtfully develop and support ideas.

"Organization" relates to the presence of a thesis, and the relationship between the ideas in the essay and this thesis. To meet the "Excellent" category, the student must write an interesting and meaningful introduction, must establish a personal focus in the writing, must control the discussion, and must write a conclusion which relates thoughtfully to the thesis. In contrast, the "Deficient" category says the "introduction relates in a limited way to the rest of the essay," that a thesis is lacking or not maintained, that ideas are not clearly developed and that the conclusion is not functional.

"Matters of choice" refers to the use of words and sentences. To receive a high mark for this category, students must make deliberate choices about words and sentences to achieve a particular purpose. The writing is precise and fluent. The deficient rating is for the use of language which is frequently inaccurate, ineffective and inappropriate.

The fourth component of the scoring guide is "matters of convention." This relates to the conventions of language such as spelling, grammar and punctuation. The excellent category is for writing which is essentially free from errors, while the deficient category is for writing

which has frequent errors.

The "thought and detail" category is worth ten marks, while the remaining three categories are worth five marks each. This is because "thought and detail" assesses the level of the student's insight and understanding of the literature in question. Alberta Education uses a similar ratio in its evaluation of English 30 Diploma Examinations.

Another part of the evaluation procedure included marking annotations about the writing in the margins of each essay, and writing summary comments on the title page. When essays were returned, students could also discuss their work individually with me.

#### Analysis of Pre-survey

The pre-survey (Appendix B) contained six statements for which students had to indicate whether they strongly disagreed, disagreed, were neutral, agreed or strongly agreed:

1. Teachers favor certain students when they mark assignments.
2. I have had at least one experience in the past year when I felt I was marked unfairly.
3. Submitting essays with no names on them is a good idea.
4. My teachers like certain students more than others.

5. When a student receives failing marks, he probably will continue to receive failing marks.
6. All teachers should mark some assignments which do not have names on them.

The results of the pre-survey suggest that the students were keenly interested in the project. Of the twenty-four students, nineteen wrote comments at the bottom of the survey.

Table 1 indicates the responses for the pre-survey. Statements one and four in the survey relate to teacher-student relationships. Sixteen students agreed or strongly agreed that teachers favor certain student when they mark assignments. Eight were neutral and only two disagreed. For statement #4, twenty-one students agreed or strongly agreed that their teachers liked certain students more than others. No students disagreed or strongly disagreed with the statement.

For the second statement, eighteen students agreed or strongly agreed that they had had at least one experience when they felt they were marked unfairly. Only one student disagreed with the statement; this student obviously was happy with the assessment of his school assignments. The students obviously liked the idea of handing in assignments with no names, because twenty-three agreed or strongly agreed with statement #3. The remaining student

Table 1  
Responses to Pre-survey

---

Question	SD	D	N	A	SA
1	0	2	6	12	4
2	0	1	5	12	6
3	0	0	1	12	11
4	0	0	3	15	6
5	1	13	4	6	0
6	1	2	1	11	9

---

was neutral.

The students did not think that once a student received failing marks, that he would continue to receive failing marks. Fourteen students disagreed or strongly disagreed with this statement. Four were neutral and six agreed.

The results for the last statement show that most students like the idea of teachers marking some assignments which do not have names on them. Twenty agreed or strongly agreed with this statement.

Of the nineteen students who wrote comments on the survey, eleven addressed the issue of favoritism. These comments confirmed the responses to the first question. A

typical comment was "Many times people are marked unfairly because once they get on a teacher's bad side they are doomed." Another said "some teachers favor the academics." In one student's mind, the reasons for favoritism included such things as being "more involved in that certain teacher's activities." One student was quite confident that when students received failing marks, it was because the work was deficient: "...if a student receives failing marks, it should indicate the student's lack of ability in that class...". This student meant that poor marks were a product of low ability. This same student said that marking unnamed essays could lead to "reverse bias" because teachers want to mark in such a way that the marks would approximate the provincial average for the Diploma Examinations. The student's concern was that teachers were, in fact, influenced by looking at students' names when they marked essays, and that they intentionally gave marks which would approximate their expectations for how the student would score on the Diploma Exam. Another student commented that "handing assignments in with no names is an incredibly good idea."

One student lamented the idea of not attaching names to assignments. He had discovered that it is important to develop positive teacher-student relationships. This strategy, he said, becomes useless when names are not put

on the essays. Evaluation, to him, meant more than scoring the actual essay. The grade was partly a product of the quality of the relationship between teacher and student. This comment relates closely to the findings of Archer and McCarthy (1988) which were cited earlier in this report. That is, teachers are influenced by a knowledge of their students and by the relationship they have with their students.

#### Analysis of Scores

Table 2 indicates the scores for each essay. Essays 1, 3 and 5 were submitted without names, while the remaining ones did have names. Several trends are evident. First, scores for each student tend to remain fairly constant. Students who generally wrote "A" (80-100%) essays received good marks whether or not their identities were known. For example, this trend held for students 1, 7, 10, 15, 18, 19, 22 and 23. Students who tended to write "B" essays received marks primarily in the "B" range (65-79%). This trend held for students 4, 5, 8, 11 and 14. Students 3, 6, 9, 16, 17 and 25 generally received scores in the "C" range (50-65%).

Students 12, 13 and 20 failed (with scores of less than 50%) almost all the essays. These students may have benefited from having their names on their work. Student 12 received scores of 25%, 32% and 48% for essays with no

Table 2  
Scores in % for each essay

student	*E1	E2	*E3	E4	*E5	E6
1	84	96	84	84	76	84
2	44	76	44	64	24	24
3	52	48	44	52	44	48
4	80	72	68	64	80	56
5	76	92	72	72	72	64
6	60	52	56	60	64	52
7	84	80	80	92	84	100
8	64	56	64	76	68	64
9	52	72	60	56	76	56
10	64	88	88	84	76	92
11	76	64	56	68	68	68
12	25	52	32	52	48	36
13	48	52	40	56	32	36
14	72	76	84	52	72	64
15	88	84	72	80	76	68
16	72	64	64	48	68	52
17	60	64	52	56	56	56
18	80	84	72	80	72	96
19	80	92	80	72	72	96
20	36	40	44	60	44	44
21	48	64	52	48	40	60
22	64	88	60	76	80	84
23	60	92	76	92	76	80
24	64	68	64	72	32	44
25	64	68	64	64	48	72

\* essays without names

name, and scores of 52%, 52% and 36% for essays with the name. A similar situation exists for Student 13. This student received scores of 48%, 40% and 32% for essays with no name, and scores of 52%, 56% and 36% for essays with the name. Student 20 received scores of 36%, 44% and 44%

for essays with a name, and scores of 40%, 60% and 44% for essays with a name.

Two students had no trend in their marks. Student 21 failed three essays (with scores of 48%, 48% and 40%), and passed three essays (with scores of 64%, 52% and 60%). Of the three failing essays, two had no name and one did have a name. Of the three passing essays, one had no names and two did have a name. Student 24 passed the first four essays, but failed the remaining two. This student admitted to not really working on these two compositions.

Table 3 indicates the means of the three essays submitted without names and the means of the three essays submitted with names. For fifteen students, the means of the two groups of essays had a difference of less than 10%. For the remaining ten students, the difference in the means of the two groups of essays was between 11-20%. Most of the students in this category had a difference between the means of 12-14%. The most serious differences occurred for students 2 and 23. The difference in means for student 2 was 17.3%. For student 23, the difference was 18%. In both cases, the mean was higher for the essays with names on them.

A t-test for nonindependent samples was conducted to analyze the difference between the scores on the essays with names and the essays without names. Table 4 provides

Table 3

Means of essays without names  
and with names

---

Student	Mean of Essays 1,3,5 Without Names	Mean of Essays 2,4,6 With Names	Difference
1	81.3	88.0	6.7
2	37.3	54.6	17.3
3	46.6	49.3	2.7
4	76.0	64.0	-12.0
5	73.3	76.0	2.7
6	60.0	54.6	-5.4
7	82.7	90.6	7.9
8	65.3	65.3	0.0
9	62.7	61.3	-1.4
10	76.0	88.0	12.0
11	66.7	66.7	0.0
12	35.0	46.7	11.7
13	40.0	48.0	8.0
14	76.0	64.0	-12.0
15	78.7	77.3	-1.4
16	68.0	54.7	-13.3
17	56.0	58.7	2.7
18	74.7	86.7	12.0
19	77.3	86.7	9.4
20	41.3	48.0	6.7
21	46.7	57.3	10.6
22	68.0	82.6	14.6
23	70.7	88.0	17.3
24	53.3	61.3	8.3
25	58.7	68.0	9.3

---

the data for this analysis. The results demonstrate that the essays scored with names attached were marked more favorably than those scored without names. The mean for essays with names was 67.45%, while the mean for essays

without names was 62.89%. The difference in means is 4.56. At the .05 level of confidence, the difference in means is statistically significant.

Table 4

t test of significance of difference among mean scores of the two marking strategies (with names and without names)

With names		Without names		D	t	prob
Mean	SD <sub>1</sub>	M2	SD <sub>2</sub>			
67.45	14.50	62.89	14.43	4.56	*2.63	2.06

\*significant at .05 level

To check for the presence of a regression toward the mean, scores for the top seven students and the bottom seven students were examined. These students were identified by choosing the seven highest scores for essays with names on them. These scores were compared to the scores for essays without names for the respective seven students. Table 5 indicates that for each of the top seven students, marks were higher when I knew their names.

Scores were also examined for the bottom seven students, who were also identified on the basis of scores for essays with names on them. Table 6 indicates that marks dropped for five of the seven students when I did not know

Table 5

Scores for essays with names and without names  
for the top seven students ranked according  
to scores for essays with names

---

Student	Mean of Essays 1,3,5 With Names	Mean of Essays 2,4,6 Without Names	Difference
7	90.6	82.7	-7.9
1	88.0	81.3	-6.7
10	88.0	76.0	-7.9
23	88.0	70.7	-17.3
18	86.7	74.7	-12.0
19	86.7	77.3	-9.4
22	82.6	68.0	-14.6

---

identities. There is no regression toward the mean. There is, however, a tendency for the marking to be more rigorous when identities are unknown.

#### Analysis of Post-survey

The post-survey contained the same six statements as the pre-survey, and again, students had to indicate whether they strongly disagreed, disagreed, were neutral, agreed or strongly agreed:

1. Teachers favor certain students when they mark assignments.
2. I have had at least one experience in the past year when I felt I was marked unfairly.
3. Submitting essays with no names on them is a good idea.

Table 6

Scores for essays with names and without names  
for the bottom seven students ranked according  
to scores for essays with names

Student	Mean of Essays 1,3,5 With Names	Mean of Essays 2,4,6 Without Names	Difference
12	46.7	35.0	-11.7
13	48.0	40.0	-8.0
20	48.0	41.3	-6.7
3	49.3	46.6	-2.7
2	54.6	37.3	-17.3
6	54.6	60.0	5.4
16	54.7	68.0	13.3

4. My teachers like certain students more than others.
5. When a student receives failing marks, he probably will continue to receive failing marks.
6. All teachers should mark some assignments which do not have names on them.

Table 7 indicates general agreement with all the statements on the post-survey except for the fifth one. Twenty students agreed or strongly agreed with the first statement; that is, that teachers favor certain students when assignments are marked.

For the second statement, results were not as conclusive. Seventeen students agreed or strongly agreed that they have had at least one experience in the past year where they felt they were marked unfairly. Four students

Table 7

Responses to Post-survey					
Question	SD	D	N	A	SA
1	0	1	4	15	5
2	1	4	3	11	6
3	0	0	1	12	12
4	0	1	1	16	7
5	1	12	5	6	1
6	0	1	2	9	13

disagreed with the statement and three were neutral.

Students particularly agreed with the third statement, which said that submitting assignments without names is a good idea. Twenty-four students agreed or strongly agreed, and one was neutral. Obviously the students like the idea of not putting their names on essays. This preference accounts for the students' response to statements one, two, three, four and six.

The fourth statement, which said that teachers liked certain students more than others, served as a kind of check for the first one. There is a difference, obviously, between liking a person, and favoring that person, but the two often go together. The students feel that this is often the case in the classroom.

The fifth statement said that if a student receives

failing marks, he will continue to receive failing marks. There was mixed reaction to this statement. Only seven students agreed or strongly agreed. Five were neutral and thirteen disagreed or strongly disagreed. It is encouraging to note the reasonably low level of agreement. This suggests that students do not feel that they are trapped and labelled by receiving low marks.

Twenty-two students agreed or strongly agreed with the six statement, that all teachers should mark some assignments with no names on them.

A comparison of Table 1 and Table 2 indicates that little change occurred regarding students' feelings and ideas about evaluation. However, for the first statement, four additional students marked the "Agree" or "Strongly Agree" boxes in the post-survey. Three of these include students 7, 8 and 15. The fourth one cannot be identified because two forms which indicated movement between "Neutral" and "Agree" or "Strongly Agree" had no names. This was not discovered until the close of the semester, because the forms were stored in a sealed envelope during the study.

Student 7 was a very capable, independent girl who consistently wrote superlative essays. Her marks were higher when I knew her identify, so perhaps her change on the post-survey was an honest reaction to this situation.

Student 8, a girl, was about six months older than the other students. Her ideas seemed to be more sophisticated than those of the other students. Her writing skills were only average, and her mark for the "matters of convention" category often were at "Deficient" or "Satisfactory." She was neutral for the first statement on the pre-survey and strongly agreed with this statement on the post-survey. Further, on the post-survey, she said "I think all assignments should remain nameless until the marking is done." She may have suspected that I was overreacting to her weak writing skills. When I knew her identity, her marks were slightly lower for two essays, but significantly higher for one essay. Student 15 changed from "Neutral" on the pre-survey to "Agree" on the post-survey for the first statement. This student, also a girl, suffered some stress during the semester due to domestic problems and the pressures of working part-time. Her marks did not seem to suffer. She did not comment on the post-survey, and I am not sure why she changed her view about teachers favoring certain students.

The post-survey showed some movement for the second statement as well. This statement related to being marked unfairly. Three additional students disagreed with the statement, so apparently they must have increased their confidence in how they were being evaluated. The students

who changed were students 19, 23 and 24. Students 19 and 23 were strong academically, while student 24 was average. On the post-survey, student 19 said "just because a teacher gives a high or low mark to a student doesn't mean that the teacher is expressing a bias, it could be because that is what the student deserves." This student valued hard work, and firmly believed that one's marks were a result of one's efforts. It is interesting that although student 24 suffered near the end of the semester - her marks dropped from the 60-70% range to the 30-40% range - she did not feel that she had been marked unfairly. She attributed her low marks to a lack of time for school assignments because of her part-time job. For the remaining four statements, there appears to be little change between the surveys.

Three other students also commented on the post-survey. Student 10 said "it's natural for certain people to prefer other people, but this bias need not appear in marking, and it's possible that a student who receives a failing mark keeps receiving failing marks simply because that's their ability." This view is similar to that of student 19, as noted in the previous paragraph. Student 5 said "when teachers know that are being tested they are more careful to present themselves as an "honest" teacher, and not mark as they truly want to." This is a

very interesting comment. First, this student seems to be a bit suspicious about the validity of the study, and thinks that I will be marking differently during the study than I would be otherwise. However, her comment does not take into account the fact that three of the six essays were scored without knowing names, and if I was indeed being "honest," I had no idea about who I was being honest with. More importantly, her second comment reveals her belief that bias is a fact of life in schools. Her comment seems to suggest that teachers have a "hidden agenda" which comes into play during evaluation. The literature review supports her claim. It is often true, that when teachers see the name, they see the mark. Some teachers may predict year-end scores and then compare their predictions with actual marks. Often, they are not far off. These expectations may bias the evaluation.

The final student to comment on the post-survey was student 17. She said "teachers should try to not only mark fairly, but give the student a fair chance to raise marks." This comment is not surprising. It comes from a student who was irresponsible and skipped class occasionally. When she returned, she would ignore missed work for weeks and then suddenly want to make it up. She felt that she should be allowed to raise her marks just before report card time, but this was not permitted for two rea-

sons. First, I did not have the time, and second, it would only encourage poor work habits.

## Findings and Conclusions

### Analytical Comments

The pre-survey and post-survey suggest that students generally feel that favoritism is a reality in the classroom. Related to this finding is the preference the students expressed for submitting anonymous assignments for evaluation. The students saw this strategy as a liberating activity which enhances the evaluation process. The weaker students realized that they were not defeated before they even started an essay, and this gave them hope. The stronger students knew that they could not rely on their name to bring them a high mark.

As indicated previously, most of the students felt that they had been unfairly evaluated at least once during the past year. This comment underlines the need for teachers to be cognizant of how they assess student work, and to be sure that students are evaluated as carefully and fairly as possible.

Students' marks for writing remain reasonably constant. In the study, marks for all the students but two followed a pattern. For example, eight students wrote pa-

pers which generally were graded above 80%. Their success results from their ability to clearly express their ideas, to think critically about the literature they study, to make connections between the literature they study and life itself, to provide suitable evidence for their ideas, and finally, to interpret this evidence to establish a logical relationship between evidence and ideas.

There is a difference in student and teacher perceptions of evaluation. Even though students felt they had been evaluated unfairly at least once during the past year, the reality of the study shows that the opposite is true. They received higher marks when their names appeared on their work. This suggests a positive teacher reaction to knowing students' identities. A negative reaction would have meant the scores for essays with names would have been lower.

There is a statistically significant difference in the means of essays with names and those without names. In the study, students received higher marks when their names were known. This was the case for seventeen students. Six students received lower marks when I knew their names, and two students had the same means for essays with names and essays without names.

### Interpretive Comments

Many of the comments in the surveys relate closely to findings quoted in the literature review in this study. For example, students generally feel that teacher bias is a reality. Some of the students referred to the quality of the student-teacher relationship as one influence on evaluation. This may partly explain the students' enthusiasm for submitting essays without names. During the study, students reacted positively to this strategy; they did not have to worry about all the extraneous factors which often become part of the evaluation process. They were confident that their evaluation would be fair.

The problem posed at the beginning of this study was: Do students' scores on essays change significantly when their identities are removed from the essays during the evaluation process? The answer to the question is "Yes." To the chagrin of the students, though, the marks will usually drop. For the sample used in the study, the marks dropped an average of 4.56%. The main reasons for the drop, for this writer, is that removing the identities and all the concomitant influences actually changes the nature of the evaluation. The pressure of meeting students' and parents' expectations is removed. The teachers' like or dislike for particular students is no longer a factor. Prior academic performance has no bearing on present

evaluation.

The study indicates that students, as well, feel more comfortable with removing their names from assignments and with standardizing the appearance of their assignments. This strategy puts all students on an equal footing, and this is one of the very important aspects of the study. Bias in student evaluation is a form of discrimination, and therefore teachers must carefully consider how they evaluate students.

Most teachers feel that evaluation is one of the worst aspects of the profession. It certainly has the potential to harm the teacher-student relationship. At the beginning of a semester, both teacher and students tackle the coming program of studies with anticipation and enthusiasm. However, as the assignments are marked, and some students begin to receive failing marks, or marks which are lower than expected, the enthusiasm wanes. No one gets excited about failing or receiving marks which are perceived as being too low. It is really frustrating for both teacher and student when a student spends hours and hours working on an assignment and still receives a mark which is lower than what the student anticipated. This situation points to the difference between the student's hope that the mark will be adequate and the reality of the finished composition. Teachers often cannot tell how long a student has

worked on an essay. The effort put into an assignment is not evaluated. Again, we must return to the theme of the study: even though the evaluation of student writing is subjective, every effort must be made to enhance fairness.

One of the ways to do this is, as Jolly (1985) and Chase (1968) suggest, to use a scoring guide. The use of a guide helps focus the evaluation on four components of the composition. The students know how they will be evaluated before they even begin writing, and further, they know the weighting for each category. The key also has another advantage: when students need to know how to improve their writing, they can look at the key to find both their strong and weak areas. The study indicated, however, that even with the use of a scoring guide, bias still existed.

The study has certain limitations. Scores may vary when names are removed and the appearance of assignments is standardized. However, variables related to the students must also be considered. Kincaid (1953) found that writers' success in doing assigned writing varied from day to day and from topic to topic. The essay scores may fluctuate because of the level of interest in a given topic, or in the literary work itself. Some students may like modern plays, but not Shakespearean plays. Some may prefer short stories to poetry. Another factor relates to time.

The students in the class may not always have time to do their best work. The lack of time may stem from the pressure of other classes, or, unfortunately, from having a part-time job. Student 24 was the most obvious example of the latter statement. Low marks on her last two essays were the direct result of having a part-time job. Finally, some students will produce a superior essay simply because of the inspiration of the moment. Because of this factor, six essays were used in the study to increase the reliability of the means.

If error in evaluation is a fact of life, as White (1988) indicates, then, in this writer's opinion, it is more fair and reasonable to err in the student's favor. This occurred in the study. Scores when names were known were higher than scores when names were not known.

Engaging in a study such as this one requires a bit of nerve - the teacher is leaving himself open to whatever happens with the scoring of the essays. To realize that the "straight A" student and the "low achiever" are suddenly anonymous certainly brings a new and refreshing perspective to the process of evaluation. Admittedly, the better writers still have the advantage, because of their ease and fluency with language. As long as the teacher reacts to the recognition of good writing, this situation is acceptable. However, our goal, as professional educa-

tors, must be to provide evaluation which is "an accurate and unbiased understanding of the strengths and weaknesses of an individual as they relate to school life and should provide focus points on which to base instruction" (Chodzinski 1988).

### Summary

This study describes one attempt to provide an "accurate and unbiased" evaluation. The objective was to see what happens to students' scores for written work when their identities are removed from the work during the evaluation process. Even though removing names and handwriting skills from compositions does remove some of the possibility for bias, the study shows that bias remained. Fortunately, the bias was in favor of the students. When Diploma Examinations are scored, Alberta Education stresses that if the marker is in doubt about whether to assign a "3" or "4," on a five point scale, for example, the marker should give the benefit of the doubt to the student and award the "4." This practice reflects the reality that we can never measure the worth of a composition with 100% accuracy.

Most teachers probably think of themselves as unbiased evaluators. This writer held a similar view, in spite of a suspicion that there might be a response to some of the

influences which come with knowledge of the writer's identity when essays are scored. The study shows that bias does exist, even when identities and handwriting skills are removed from the compositions. The evaluator should at least be aware of how the characteristics of the marker, the characteristics of the writing, and the characteristics of the writers all operate to influence the assessment of the work.

## References

- Archer, J. and McCarthy, B. (1988). Personal biases in student assessment. Educational Journal, 30 (2), 142-144.
- Breland, H.N. and Jones, R.J. (1984). Perceptions of writing skills. Written Communication, 1.1 (Jan. 1984), 101-119.
- Chase, C.I. (1968). The impact of some obvious variables on essay test scores. Journal of Educational Measurement. 5 (4), 315-316.
- Chodzinski, R.T. (1988). Teacher strategies for non-biased student evaluation and program delivery: a multicultural perspective. The Canadian Modern Language Review. 45 (1), 65-75.
- Cooper, C. and Odell, L. (1977). Evaluating Writing: Describing, Measuring, Judging. Urbana: National Council of Teachers of English.
- Diederich, P.B. (1974). The effect of bias. Measuring Growth For English. Urbana: National Council of Teachers of English.
- Freedman, S.W. (1979). How characteristics of student essays influence teachers' evaluations. Journal of Educational Psychology. 71 (3), 328-338.
- Freedman, S.W. (1982). Some reasons for the grades we give compositions. English Journal. 71 (Nov. 1982), 86-89.
- Freedman, S.W. and Calfee, R.C. (1983). Holistic assessment of writing: experimental design and cognitive theory. Research For Writing: Principles and Methods. Ed. Mosenthal, Tamor, Walmsley. New York: Longman.
- Graham, S. and Dwyer, A. (1987). Effects of the learning disability label, quality of writing performance, and examiner's level of expertise on the evaluation of written products. Journal of Learning Disabilities. 20 (5), 20-46.
- Gronland, N.E. (1981). Measuring complex achievement: the essay test. Measurement and Evaluation in Teaching. New York: Macmillan Publishing Company., Inc.
- Harris, W.H. (1977). Teacher response to student writing: a study of the response patterns of high school English teachers to determine the basis for teacher judgement of student writing. Research in Teaching English. Fall, 1977, 175-85.

- Jolly, P.B. (1985). Evaluating the evaluator. Paper presented at the Annual Meeting of the Conference on College Composition and Communication, Minneapolis, MN, March 21-23, 1985.
- Marshall, J.C. and Powers, J.M. (1969). Writing neatness, composition errors, and essay grades. Journal of Educational Measurement. 6 (2), 97-98.
- McCaig, R.A. (1982). How to (and How Not to Conduct an Assessment of Student Writing. (unpublished report, Grosse Pointe Public School System).
- Nyberg, A. (1988). Points to ponder when scoring high school English. Alberta English 27 (1), 9-11.
- Peacock, M. (1988). Does the "look" of word processed text influence marks? Journal of Computer Assisted Learning, 4 162-172.
- Romano, T. (1987). Clearing the Way: Working With Teenage Writers. Portsmouth: Heineman Educational Books, Inc.
- Rosenblatt, L. (1978). The Reader, the Text and the Poem. (1978). Carbondale: Southern Illinois University Press.
- Sloan, C.A. and McGinnis, I. (1978). The Effect of Handwriting on Teachers' Grading of High School Essays.
- Wade, B. (1978). Responses to written work: the possibilities of utilizing pupils' perceptions. Educational Review 30 (June 1978), 149-58.

Appendix A

READING LIST  
MR. FISHER - ENGLISH 30

Bunyan - The Pilgrim's Progress  
Callaghan - More Joy in Heaven  
Cormier - I Am the Cheese  
Cormier - The Bumblebee Flies Anyway  
Cormier - The Chocolate War  
Cormier - Beyond the Chocolate War  
Dickens - Great Expectations  
Dreiser - An American Tragedy  
Findlay - The Wars  
Grove - Settlers of the Marsh  
Hardy - Tess of the D'urbervilles  
Hawthorne - The Scarlet Letter  
Homer - The Odyssey  
Kesey - One Flew Over the Cuckoo's Nest  
Lawrence - The Stone Angel  
Orwell - Nineteen Eighty-four  
Rand - The Fountainhead  
Richler - The Apprenticeship of Duddy Kravitz  
Robbins - A Stone For Danny Fisher  
Steinbeck - East of Eden  
Steinbeck - The Grapes of Wrath  
Tolkien - The Hobbit  
Twain - The Adventures of Huckleberry Finn

Appendix B

EVALUATION OF SCHOOL ASSIGNMENTS

pre-survey

Name: \_\_\_\_\_

Students: This survey is part of the research project you will be involved in for the next twelve weeks. The purpose of this survey is to collect information about your feelings and ideas about how you are evaluated in school.

Please mark the appropriate column for each statement.

SD Strongly Disagree  
D Disagree  
N Neutral  
A Agree  
SA Strongly Agree

SD D N A SA

1. Teachers favor certain students when they mark assignments.
2. I have had at least one experience in the past year when I felt I was marked unfairly.
3. Submitting essays with no names on them is a good idea.
4. My teachers like certain students more than others.
5. When a student receives failing marks, he probably will continue to receive failing marks.
6. All teachers should mark some assignments which do not have names on them.

Please respond in writing to any of the above statements, or about any aspect of this research project.

-----  
-----  
-----  
-----  
-----  
-----  
-----  
-----

Appendix C

EVALUATION OF SCHOOL ASSIGNMENTS

post-survey

Name: \_\_\_\_\_

Students: This survey is part of the research project you will be involved in for the next twelve weeks. The purpose of this survey is to collect information about your feelings and ideas about how you are evaluated in school.

Please mark the appropriate column for each statement.

- SD Strongly Disagree
- D Disagree
- N Neutral
- A Agree
- SA Strongly Agree

SD    D    N    A    SA

1. Teachers favor certain students when they mark assignments.
2. I have had at least one experience in the past year when I felt I was marked unfairly.
3. Submitting essays with no names on them is a good idea.
4. My teachers like certain students more than others.
5. When a student receives failing marks, he probably will continue to receive failing marks.
6. All teachers should mark some assignments which do not have names on them.

Please respond in writing to any of the above statements, or about any aspect of this research project.

-----  
-----  
-----  
-----  
-----  
-----  
-----

Appendix D

Dear Parent:

I am conducting a study of the assessment of student writing. The purpose of the study is to determine if there is teacher bias in the assessment of essays written by students. I anticipate your child and others will benefit from participation in this study by developing and reinforcing their word processing skills and by becoming more sensitive to the issues related to and the processes involved in assessing essays. I would like your permission for your son or daughter to participate in this study.

As part of this research, your son or daughter will be asked to submit every second essay without a name on it. These essays are part of the regular course work; no additional writing is involved. Students will place a number on these essays for later identification. Also, during the study, each participant will be using computers at school to write the essays.

Please note that all information will be handled in a confidential and professional manner. All names, locations and any other identifying information will not be included in any discussion of the results. You also have the right to withdraw your child from the study without prejudice at any time.

If you choose to do so, please indicate your willingness to allow your child to participate by signing this letter in the space provided below, and return the letter to the school with the student.

I very much appreciate your assistance in this study. If you have any questions please feel free to call me at school at 527-3371 or at home at 527-0282. Also feel free to contact any member of the Faculty of Education Human Subjects Research Committee at The University of Lethbridge if you wish additional information. The chairperson of the committee is Nancy Grieg.

Yours sincerely,

Keith Fisher  
Teacher - Medicine Hat High School (527-3371)

(Please detach and forward the signed portion.)

-----  
Bias In Assessment: Product or Writer?

I agree to allow my child, \_\_\_\_\_, to participate in this study.

\_\_\_\_\_  
Name

\_\_\_\_\_  
Signature

\_\_\_\_\_  
Date

**ENGLISH 20 AND 30  
CRITICAL RESPONSE SCORING GUIDE**

**THOUGHT AND DETAIL**

<b>EXCELLENT:</b> Insightful ideas are supported by carefully chosen details. Literary interpretations are perceptive and defensible.	<b>10, 9, 8</b>
<b>PROFICIENT:</b> Insightful ideas are supported by appropriate ideas, OR conventional ideas are supported by carefully chosen details. Literary interpretations are thoughtful and defensible.	<b>7</b>
<b>SATISFACTORY:</b> Ordinary ideas are supported by appropriate details. Literary interpretations are defensible.	<b>6, 5</b>
<b>DEFICIENT:</b> Ordinary ideas are weakly supported, or are accompanied by inappropriate details. Literary interpretations are incomplete or superficial.	<b>4, 2</b>

**ORGANIZATION**

<b>EXCELLENT:</b> The introduction provides direction for the reader and/or provokes further reading. A personal focus is established. The controlling ideas is successfully sustained and developed in a clear manner. The conclusion relates thoughtfully to the thesis.	<b>5</b>
<b>PROFICIENT:</b> The introduction provides direction for the reader. The controlling idea is focused and is generally sustained. The development of the controlling idea is clear and generally coherent. The conclusion relates appropriately to the thesis.	<b>4</b>
<b>SATISFACTORY:</b> The introduction provides a general direction for the reader. The thesis provides a focus at the beginning that is mechanically maintained. The development of the thesis is clear, but coherence occasionally falters. The conclusion relates functionally to the thesis.	<b>3</b>
<b>DEFICIENT:</b> The introduction relates in a limited way to the rest of the essay. A thesis is lacking or is not maintained in the development of the essay. The ideas are not clearly developed. The conclusion is not functional.	<b>2, 1</b>

**MATTERS OF CHOICE**

<b>EXCELLENT:</b> Choices appear to have been made deliberately to achieve a particular purpose. The selection and use of words and sentence structures is effective and polished. Diction is effective and specific. The writing is precise and fluent.	<b>5</b>
<b>PROFICIENT:</b> Choices frequently appear to have been made deliberately to achieve a particular purpose. The selection and use of words and sentence structures is generally effective. Diction is appropriate. The writing is clear and fluent.	<b>4</b>
<b>SATISFACTORY:</b> Choices occasionally appear to have been made deliberately to achieve a particular purpose. The selection and use of words and sentence structures is generally clear. Diction is adequate but may be lacking in specificity.	<b>3</b>
<b>DEFICIENT:</b> Choices do not appear to have been made deliberately to achieve a particular purpose. The selection and use of words and sentence structures is frequently inaccurate and ineffective. Diction is frequently inaccurate and/or inappropriate.	<b>2, 1</b>

**MATTERS OF CONVENTION**

<b>EXCELLENT:</b> This writing is free from errors in spelling, punctuation and grammar.	<b>5</b>
<b>PROFICIENT:</b> This writing is essentially free from errors in spelling, punctuation and grammar.	<b>4</b>
<b>SATISFACTORY:</b> This writing has occasional errors in spelling, punctuation and grammar.	<b>3</b>
<b>DEFICIENT:</b> This writing has frequent errors in spelling, punctuation and grammar.	<b>2, 1</b>