CHAPTER 1: INTRODUCTION

Introduction

How can teachers determine whether a student's piece of writing meets the standard of exceptional versus that of good? How can a child's science project, which may involve drawing, writing, diorama, oral presentation, and other elements, be accurately evaluated? How can slight, gradual progress in the development of speaking skills be measured? One increasingly popular method receiving high marks across the country are rubrics, a type of scoring guide used to assess instruction and performance according to pre-determined expectations and criteria (Martin-Kniep, 1998). Since rubrics set forth such specific criteria and define precise requirements for meeting those criteria, each learner is aware of and understands exactly what is required to achieve a specific score on each task.

Jackson and Larson (2002) found that rubrics can be used as both formative and summative assessment tools and can be implemented in three primary ways: first, they can be given and discussed prior to assignment completion as a means for students to evaluate their performance during the work phase; secondly, teachers may engage students in the rubric development, providing students an opportunity to develop assessment criteria and take ownership in their learning; and finally teachers can use rubrics to award a final grade after the assignment is submitted. The goal of this research study was to assess the impact of these different methods of rubric implementation on achievement and student attitude.

Statement of the Problem

The purpose of this study was to assess the impact of different methods of rubric implementation on achievement and student attitude.

Research Question

What was the impact of using different methods of rubric implementation on achievement and student attitude?

Sub Questions

1. How did different rubric implementation techniques affect achievement?
2. How did different rubric implementation techniques affect achievement with regard to gender?
3. How did different rubric implementation techniques affect achievement of regular education students?
4. How did different rubric implementation techniques affect achievement of special education students?
5. How did different rubric implementation techniques affect attitude with regard to gender?
6. How did different rubric implementation techniques affect attitude of regular education students?
7. How did different rubric implementation techniques affect attitude of special education students?

Null Hypothesis #1

Different methods of rubric implementation have no effect on achievement.

Research Hypothesis #1

Different methods of rubric implementation have an effect on achievement.

Null Hypothesis #2

Different methods of rubric implementation have no effect on achievement with regard to gender.

Research Hypothesis #2

Different methods of rubric implementation have an effect on achievement with regard to gender.

Null Hypothesis #3

Different methods of rubric implementation have no effect on achievement of regular education students.

Research Hypothesis #3

Different methods of rubric implementation have an effect on achievement of regular education students.

Null Hypothesis #4

Different methods of rubric implementation have no effect on achievement of special education students.

Research Hypothesis #4

Different methods of rubric implementation have an effect on achievement of special education students.

Null Hypothesis #5

Different methods of rubric implementation have no effect on student attitude.

Research Hypothesis #5

Different methods of rubric implementation have an effect on student attitude.

Null Hypothesis #6

Different methods of rubric implementation have no effect on attitude with regard to gender.

Research Hypothesis #6

Different methods of rubric implementation have an effect on attitude with regard to gender.

Null Hypothesis #7

Different methods of rubric implementation have no effect on attitude of regular education students.

Research Hypothesis #7

Different methods of rubric implementation have an effect on attitude of regular education students.

Null Hypothesis #8

Different methods of rubric implementation have no effect on attitude of special education students.

Research Hypothesis #8

Different methods of rubric implementation have an effect on attitude of special education students.

Definitions of Terms

For the purpose of this study, the following terms were defined.

Rubric was defined as a scoring guide used in assessing instruction and performance according to pre-determined expectations and criteria (Martin-Kniep, 1998).

Achievement was defined as a measurable increase in the understanding of and the ability to perform an academic task.

Attitude was defined as a student's feelings or beliefs regarding a specific topic or subject.

Rubric implementation was defined as the method in which a rubric is used for assessment purposes.

Regular education is defined as a plan/program for students whose ability meets the instructional demands of the classroom and can achieve commensurate with his or her age and ability level.

Special education was defined as a program with specialized instruction for students whose ability to meet instructional demands of the classroom and to achieve commensurate with his or her age or ability level is severely delayed in any of the following areas: oral expression, listening comprehension, written expression, basic reading skills, reading comprehension, mathematical calculation, mathematical reasoning, and intellectual ability.

Limitations

1. The study sample was chosen for convenience and was not random.
2. It was assumed that a survey created by the researcher was an accurate reflection of student attitudes.
3. The subjects were the students of the researcher.
4. Conclusions based upon data from this study were restricted to the sample and may not have been applicable to other classrooms.

CHAPTER 2: REVIEW OF RELATED LITERATURE

Introduction

The purpose of this study was to assess the impact of different methods of rubric implementation on achievement and student attitude. Although the researcher did not find any studies that exactly replicated this study, the following is a summary of the related professional literature and a synthesis of its pertinence to this research. Following a topical pattern, the chapter begins with the reliability and validity of rubrics and raters. Research using rubrics in assessment and teaching of regular education students is discussed next, followed by using rubrics in assessment and teaching of students with special needs. The chapter concludes with a summary of the major findings from the reviewed research.

Reliability and Validity of Rubrics and Raters

During the 1996-1997 school year, thirty-seven of forty-seven states used some form of performance assessment in their testing programs (Johnson, Penny, and Johnson, 1998). These open-ended items were usually scored with a holistic or analytic rubric with four to six-point scales, providing a framework for the scoring process "designed to minimize measurement errors resulting from rater judgment in scoring essays and similar open-ended assessments" (p. 269). Even with this framework in place, there appeared to be some concern in whether different examiners who evaluated student responses (using a rubric) drew similar conclusions about a student's performance. The following research studies illustrated this point.

Stuhlmann, Daniel, Dellinger, Kenton, and Powers (1999), professors at Louisiana State University, investigated whether training raters to interpret the scoring dimensions on a rubric would increase reliability and rater consistency when scoring writing samples. The study examined forty kindergarten and first grade teachers' abilities to use an established rubric to rate

twenty, first-grade writing samples. Twenty-three of the teachers were trained to interpret the scoring dimensions of the rubric, while seventeen were not. The trained teachers attended an hour long training session in which each section of the rubric was explained, were provided with specific examples (anchor papers) of what was acceptable at each level of the rubric, and received practice in scoring. The untrained raters were told to use the established rubric to rate each writing sample and make their own judgments regarding scoring choices. The results of the study indicated that since the trained group was more uniform in their interpretation of the scoring categories, and there was adequate agreement among raters. Training increased teachers' abilities to consistently apply the scoring criteria found in the rubric thus deeming the overall scores more reliable.

Schafer (2001) found comparable results. In a similar study, he examined teachers' knowledge of assessment rubrics on the achievement of high school students in the four subject areas of algebra, biology, English, and government. One member of each of the forty-six participating teacher pairs received extensive rubric training through a two-day workshop, while the other only received a basic introduction to the scoring rubric. After the tests were administered and then evaluated by the pairs using the generic rubric, Schafer concluded that there was support for instructional value in teacher knowledge of rubrics in general. The strength was most significant, he noted, in the content areas of biology and algebra.

Penny, Johnson, and Gordon (2000), researchers for the Center of Creative Leadership in North Carolina, took a different approach to attempt improving inter-rater reliability in the scoring of performance assessments. A study was conducted in which raters were allowed to indicate an augmented integer rating that best described the level of proficiency represented by a writing sample. This study involved 120 randomly selected writing samples from Georgia's

eleventh grade writing assessment. Two selected raters from the state's scoring program scored the writing samples using a four-point analytic rubric. The raters first each assigned an integer score to a sample that was based on the rubric criteria and benchmark papers and then were asked to indicate whether the score should be augmented (positively or negatively) or remain as is. Throughout the scoring process, raters chose to augment thirty-five percent of their ratings, with forty-three percent of those augmentations positive and fifty-seven percent negative. Penny et al. concluded that the use of rating augmentation tends to improve most instances of inter-rater reliability, "although the percentage of exact and adjacent agreement decreases because of the increased number of rating possibilities" (p. 270).

The assessment of post-secondary academic skills today can be comprehensive in nature. Using a rubric versus a more traditional scoring method may provide more meaningful and stable assessment results. "The stability of assessment results, however, rests on the scales ability to lead to a common and uniform interpretation of student performance" (Simon and Forgette-Giroux, 2001, p. 5). Simon and Forgette-Giroux of the University of Iowa studied inter-rater and intra-rater aspects of reliability with a qualitative, descriptive rubric used assess college students' academic competencies. They found that inter-rater and intra-rater aspects of reliability were greatly improved by attaching the actual scoring rubric to a course outline and by illustrating its different components and use early in a course. First, examining the rubric enabled both the graduate and undergraduate students to better self assess their portfolios while tracking progress of their targeted skill (i.e. ability to analyze research studies). Second, the rubrics provided qualitative information regarding a performance in relation to a more desired one. Finally, it presented various performance levels with specific criteria to help students develop targeted

academic skills. Simon and Forgette-Giroux concluded that providing a rubric led to firm assessment results and that it can be very useful in higher education.

In 1997, a comparative study was conducted by Gearhart, Herman and Novak of the National Center for Research on Evaluation, Standards and Student Testing to determine the technical validity of two narrative-writing assessment rubrics, an existing holistic rubric and a new, more narrative-specific framework called the Writing What You Read (WWYR) rubric. Percentages of agreement between the five raters, correlations between raters, and generalizability coefficients were examined. Gearhart et al. found that there was a pattern of greater support for the reliability of the WWYR rubric across all three applications. Based on these results, the researchers believed there was sufficient evidence that the WWYR narrative rubric could be used reliably in assessing narrative performance tasks.

Using Rubrics in Assessment and Teaching of Regular Education Students

With rubrics being such frequent topics of discussion at many state, regional and national education conferences, Wenzlaff, Fager, and Coleman (1999) decided to investigate the merits and values of rubrics and distinguish how educators were using rubrics in their teaching. Their study involved one-hundred elementary and secondary educators (from all grade levels and content areas) who were surveyed regarding feedback to students, assessing work, defining rubrics, and use of rubrics. From the survey results, the researchers found that educators believed students' abilities and interests, state standards/curriculum guides, and students' involvement should be taken into consideration when designing a rubric, the terms rubric and assessment were often used interchangeably among elementary, secondary and postsecondary educators, and rubrics could make assessment more meaningful, elucidate expectations and provide better feedback to students.

Rhoads (1998) found rubrics to be developmentally appropriate learning and assessment devices for kindergarten students. His research focused on the use of a writing rubric for self-assessment purposes and as a reliable assessment tool. In the study, Rhoads allowed his students to have ownership of the rubric by involving them in developing the criteria for good writing. Their ideas were then turned into an understandable rubric that he and the students could use to assess writing activities. Over the duration of the four-week study, Rhoads found that the kindergarten students became very confident and comfortable with the rubric. Many students strove to meet the highest level of criteria which was thumbs up writing. Self-assessing daily writing activities also became a routine to check learning and growth. Scores on the writing tasks improved for most students as well.

A study by Ceprano and Garan (1998) yielded positive results using a rubric with average ability first grade students. The semester-long action research study involved eighteen language arts university students engaged in a pen-pal project with a group of first grade students. An assessment rubric was created to chart the children's growth in voice and some of the technical aspects of writing. After the letters were rated each week with a holistic rubric, Ceprano and Garan discovered that the children's voices began to emerge more and more over the span of the study. They determined the rubric was an authentic method in charting the growth of voice and conventions of language.

Anderson and Woods (2002) conducted a research project in a ninth grade physical education class. Twelve females and seven males designed and implemented an evaluation rubric for an aerobics unit. In small groups, students designed an aerobics routine to present to the class. Rubrics were then created by each group to evaluate their aerobics routine. The researchers shared the belief of Wiggins (1998) who stated, "Sharing the responsibility for

evaluation with students gives them a degree of responsibility for their own education and encourages them to learn something about the nature of the evaluation itself" (p. 58). Input from rubrics the groups designed was then used to design a single rubric that the class agreed on and applied. As the routines were presented, the classmates, researchers and presenters scored the criteria using the adopted rubric. Anderson and Woods discovered that in evaluating students' performances, the males were much more consistent than the females. The researchers also noted that if a student became capable of effective self-evaluation, improved performance resulted.

Findings from a study conducted by Marzano (2000) showed that using rubrics as opposed to a traditional point system had a direct effect on middle school science students' learning. The science achievement of students who had been assessed using a criterion-referenced, rubrics based approach to students assessed using a norm-referenced, point based approach was compared in the study. One earth science class, the experimental group, was assessed on a lab experiment using a rubric. The other, the control group, was assessed using the regular points based approach. Marzano found the average score of the group that was evaluated using the rubric was fourteen percentage points higher than it was for the group that the points approach was used. In addition, he noted the group that used a rubric had more positive attitudes about science than did the points group.

Using Rubrics in Assessment and Teaching of Students with Special Needs

Linda Superville (2001), a teacher and researcher in the West Indies, initiated an action research project in which a rubric was used as an oral assessment measure with underachieving students. She believed that developing the students' oral dialogue would increase their written expression. The nine-week study, called Operation Talk, was implemented in her social studies

class. Twenty-nine academic underachieving students who spoke nonstandard English were studied. After processing the results of pre-study questionnaires, Superville met with the group and discussed the rubric that was going to be used to assess their oral communication skills. With a large-scale version of the scoring rubric posted in the classroom to view during the period of the study, Superville found that students quickly became familiar with what was expected of them and were easily able to weigh the value of their classmates' oral responses using the rubric criteria each week. Superville also concluded that even though developing a wide range of questions was time consuming and the students were initially self-conscious about expressing answers orally, positive results were gained. The students had taken more ownership in their learning and with the help of the rubric to peer-assess their classmates, they became improved evaluators of their classmates' performances as well.

Schirmer, Bailey, and Fitzgerald (1999) explored whether a writing assessment rubric could be used as an effective teaching strategy for students who were deaf. Their year long study, a cross-sectional design using both qualitative and quantitative data, included ten fifth and seventh grade children who were deaf. Schirmer et al. believed that a rubric strategy approach to writing could be used to highlight good qualities of writing and that the students could then reflect on those qualities. From September through June, students wrote several compositions that reflected the students' abilities to use language functionally. The compositions were evaluated with a five-point analytic rubric by graduate level students who had no prior rubric training. Quantitative analysis of compositions written early and later in the year showed that the use of the rubric as a teaching strategy considerably improved the areas of topic, content, story development and organization. It did not improve, however, the areas of text structure, voice, word choice, or mechanics. Qualitative analysis indicated that the personal functions of

expressing emotions, imaginative story telling and discussing cause/effect appeared consistently in the students' compositions. Schrimer et al. found that since the rubric strategy included aspects of both clear skills instruction and writing immersion, it was an effective instructional strategy in helping these deaf students become more proficient writers.

James and Abbott (2001) discovered that using a workshop approach to writing paired with a six-trait assessment rubric was effective in promoting the writing growth of students with learning disabilities. Every day for nine weeks, students were given a short mini-lesson focused on one of the six traits of good writing and then wrote for thirty minutes. Graphic organizers were used during pre-writing and students were encouraged to incorporate the six traits of writing into their projects. Final drafts were scored by both the teacher and one of the researchers using the six-trait writing rubric. After the nine week instruction, James and Abbott concluded that the integration of the six-trait rubric and writing workshop model greatly improved these fourth-grade students' writing abilities.

Major Findings

Current researchers seem to suggest a positive impact on writing achievement with the use of scoring rubrics in classrooms. Achievement also can increase when teachers share the evaluation rubrics with their students. Soles (2001), Wichita State University's writing program director, found that when rubrics were shared with students, "a positive correlation existed between the students' knowledge of the evaluative criteria and the grades the students receive on an essay when the evaluative criteria is applied" (p. 4). In his study of first-year college students, an experimental group was given a list of criteria upon which a writing assignment would be graded. A control group was not given the criteria. The instructor discussed the criteria with the experimental group and encouraged them to draft and revise their

essay in the context of the rubric. The instructor used the same scoring rubric to evaluate both the control group and experimental groups' essays. Soles discovered that the students who knew the criteria wrote more effectively than the students who did not. Sharing the rubric with the students, he noted, made them feel empowered and reduced writing apprehension as well.

Andrade-Goodrich (1999) also studied whether providing students with instructional rubrics had an effect on students' writing and their understanding of the qualities of good writing. Students in eighth grade classes wrote three different essays (persuasive, historical fiction, and autobiographical), one month apart. Questionnaires asking students how their teachers decided the grade for their work were given to all students after the final essays were written. Before constructing the first draft of each essay, students in the treatment classes were introduced to and given a rubric by their teachers. Students in the control classes were not given a rubric. Both the researcher and the classroom teacher scored the essays with the rubrics given to the treatment group. From analyzing the scores from both classes, Andrade-Goodrich determined instructional rubrics could help students write better. It was also determined from the questionnaires that the control students had a poorer understanding of how grades were given. The treatment group, however, mentioned high quality writing traits, similar to the traits in the rubric criteria.

Summary

The data found in this review of related literature suggested to the researcher that there was a correlation between how rubrics are implemented in classrooms and the writing achievement of both regular education and special needs students. Only one of the studies found focused on rubric use and attitudes. The related research indicated that this topic is worthy of

study, but further studies are necessary in order to more accurately determine the impact of different methods of rubric implementation on achievement and attitude.

CHAPTER 3: PROCEDURES

Research Design

The purpose of this study was to assess the impact of different methods of rubric implementation on achievement and student attitude. The design of the study was quasi-experimental, action research that produced both qualitative and quantitative data. The goal of the study was to assess which of two rubric implementation methods produced higher achievement levels and positive student attitudes.

Students in two classes were given a pre-survey (Appendix A) designed by the researcher to determine attitudes regarding assessment techniques and writing. The survey was administered the first week in December 2002 so that the results were not influenced by holiday anticipation. Following the survey, a treatment was given to one class over a two week period. The second class, the control group, did not receive a treatment. Treatment consisted of the researcher showing and discussing the analytic evaluation rubric (Appendix B) on an overhead transparency with the group of students prior to drafting the writing assignment regarding a favorite family tradition. The control group did not receive the treatment since a rubric was used to evaluate the tradition paragraphs once the final drafts were submitted. After the treatment concluded, the writing assignments were evaluated with the analytic rubric and a post attitude survey (Appendix C) was administered to all students. Data was entered into a spreadsheet, which was the framework for analysis.

Sample Description

The sample in this research study included two sixth grade English classes at Tri-County Middle School in Plainfield, Wisconsin. The sample consisted of forty-three students, thirteen males and nine females in the treatment group and ten males and eleven females in the control

group. The sixth grade students were between the ages of eleven and twelve years old at the beginning of the study. The majority of the students were Caucasian, with seven being Hispanic.

At the time of the study, Tri-County Area School District consisted of an early childhood through twelfth grade school with approximately 840 students. It was located in Waushara County in the central region of Wisconsin. Students from the surrounding counties of Adam and Portage also attended the school. At the time of the study, approximately forty-two percent of the students received free or reduced lunch.

At the time of the study, Plainfield was a small, rural farming community of approximately 900 people. Plainfield Trucking and the school itself employed the majority of the working population. Other employment opportunities in the area included potato and dairy farming, tree trimming, and factory/shift work in nearby cities. Plainfield had received a large amount of permanent migrant workers in the previous five years, initiating an increasing amount of limited language proficiency (LLP) people in the town.

Instruments

One main instrument used to obtain data for this study was a seven question attitudinal pre-survey and ten question attitudinal post-survey designed by the researcher. The survey was designed to measure student attitudes about writing and assessment techniques both before and after the treatment was given. Students could respond to the survey questions by choosing either: four (strongly agree), three (agree), two (disagree) or one (strongly disagree). Student responses to the open-ended questions were also included in the data. Another main instrument used to assess achievement was an analytic evaluation rubric also created by the researcher. Students were evaluated on the following five categories: topic/closing sentence, supporting details, conventions, tradition elements, and illustration. Students could have received either an

excellent (five points), good (four points), satisfactory (three points) or needs improvement (two points) for each category based on the criteria outlined in the evaluation rubric (Appendix B). The score received for each category was added together and then divided by twenty-five points to obtain an achievement score on the essay.

Data Analysis

The researcher analyzed and compared the essay achievement scores and results of the pre/post-survey by using a Student t-test at the .05 level of significance. The achievement scores of the tradition essay from both implementation methods was also analyzed and compared using a Student t-test at the .05 level of significance. Bar graphs were used to illustrate quantitative totals. Mean scores were determined from data collected before and after the treatment. The researcher attempted to analyze the qualitative data by looking for similarities and differences.

Calendar

This study was conducted over a two-week period during the first semester of the 2002-2003 school year. On December 4th, 2002 the students in both the treatment and control groups were given a pre-survey to complete. The researcher discussed the evaluation rubric with the treatment group on December 4th, 2002. Both the treatment and control groups drafted the tradition essay the week of December 4th, 2002. All students completed the post-survey on December 11th, 2002. Data was analyzed in December and entered into a spreadsheet. Chapters four and five were written in February 2003 and submitted during March 2003. The final approved copy was submitted prior to June 1, 2003.

Budget

The cost of conducting this study was minimal.

CHAPTER 4: RESULTS

Introduction

The purpose of this study was to assess the impact of different rubric implementation methods on achievement and student attitude. The study design was quasi-experimental action research and produced both quantitative and qualitative data. Data was collected from forty-three, sixth-grade English students using pre and post attitudinal surveys and essay achievement scores. This chapter provided a discussion of the results that were reported for the surveys and achievement scores in both quantitative and qualitative form. The results are summarized at the conclusion of this chapter.

Results

Twenty-one sixth-grade students in the control group and twenty-two sixth grade students in the treatment group wrote an essay on a favorite family tradition. The essay was then evaluated by the researcher with an analytic evaluation rubric (Appendix B). The evaluation rubric was shown on an overhead transparency and discussed with the treatment group prior to drafting the tradition essay. Each of the specific criteria (topic sentence, supporting details, conventions, tradition elements, and illustration) and the proficiency levels (excellent, good, satisfactory, needs improvement) were explained to the treatment group. The control group did not view the evaluation rubric prior to drafting the tradition essay. The scores received for each criteria of the rubric were added together and divided by twenty-five points to obtain an essay achievement score.

The graph in Figure 4.1 illustrates the essay achievement scores of the students in the control group. One student in the control group received an achievement score of ninety-six percent, seven students received a ninety-two percent, three received an eighty-eight percent, five received an eighty-four percent, three students received an eighty percent and two received a score of seventy-six percent. The average achievement score of the control group was eighty-six percent.

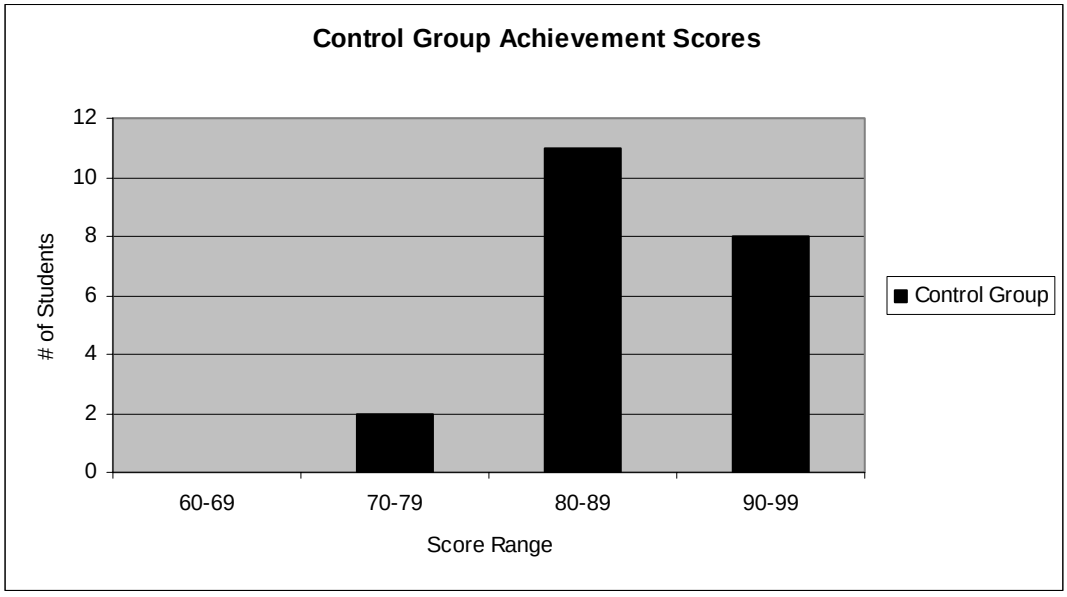

Figure 4.1

The graph in Figure 4.2 illustrates the essay achievement scores of the students in the treatment group. Two students in the treatment group received an achievement score of ninety-six percent, four students received a ninety-two percent, seven students received an eighty-eight percent, eight students received an eighty-four percent and one student received an eighty percent. The average achievement score of the treatment group was eighty-eight percent.

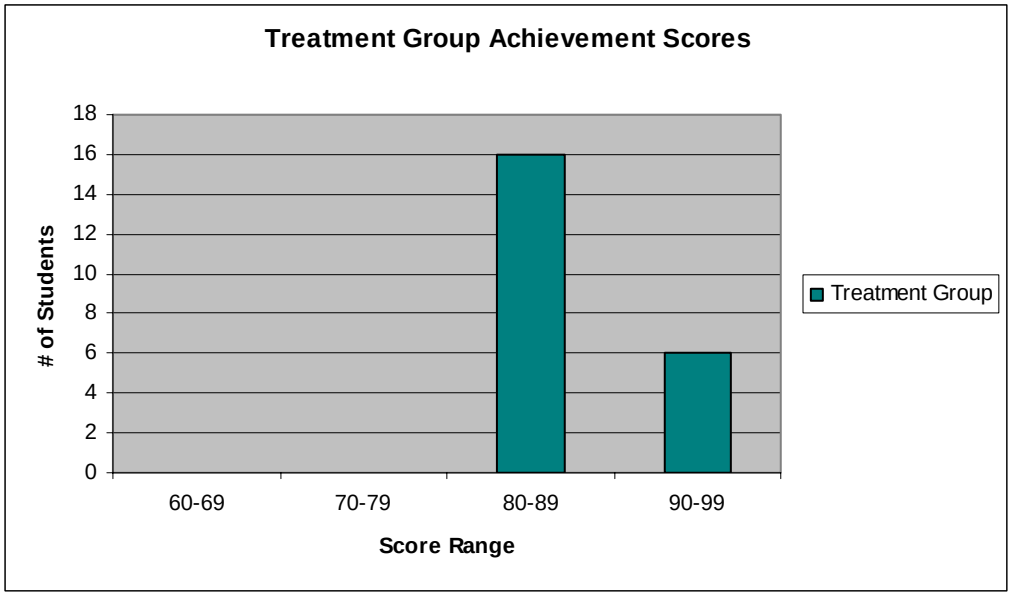

Figure 4.2

The researcher's first null hypothesis was that different methods of rubric implementation have no effect on achievement. A Student t-test at the .05 level of significance was conducted to test the hypothesis. The achievement scores of the students in the control group and the achievement scores of the students in the treatment group were compared. The calculated value of the statistic for this test was -0.74. The critical value was 2.01. The degrees of freedom were 41. From this analysis, the researcher concluded that the data was insufficient to reject the null hypothesis.

The graph in Figure 4.3 illustrates the essay achievement scores of the ten males in the control group and the thirteen males in the treatment group. In the control group, two male students received a ninety-two percent, one male received an eighty-eight, four received an eighty-four percent, two males received a score of eighty percent, and one received a seventy-six percent. The average male achievement score in the control group was eighty-four percent. In the treatment group, one male received a ninety-six percent, two received a ninety-two percent, three males received an eighty-eight percent, six received an eighty-four percent, and one received a seventy-six percent. The average male achievement score in the treatment group was eighty-six percent.

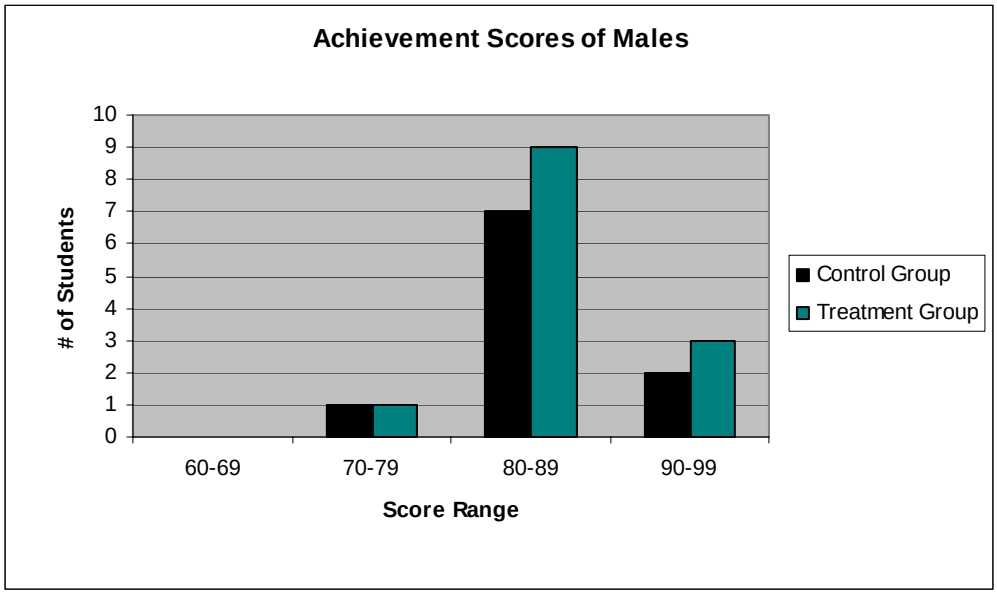

Figure 4.3

The graph in Figure 4.4 illustrates the essay achievement scores of the eleven females in the control group and the nine females in the treatment group. In the control group, one female student received a ninety-six percent, five received a ninety-two percent, two received an eighty-eight percent, one received an eighty-four percent, one received an eighty percent and one received a seventy-six percent. The average female achievement score in the control group was eighty-eight percent. In the treatment group, two received a ninety-two percent, four received an eighty-eight percent, two received an eighty-four and one received an eighty percent. The average female achievement score in the treatment group was eighty-seven percent.

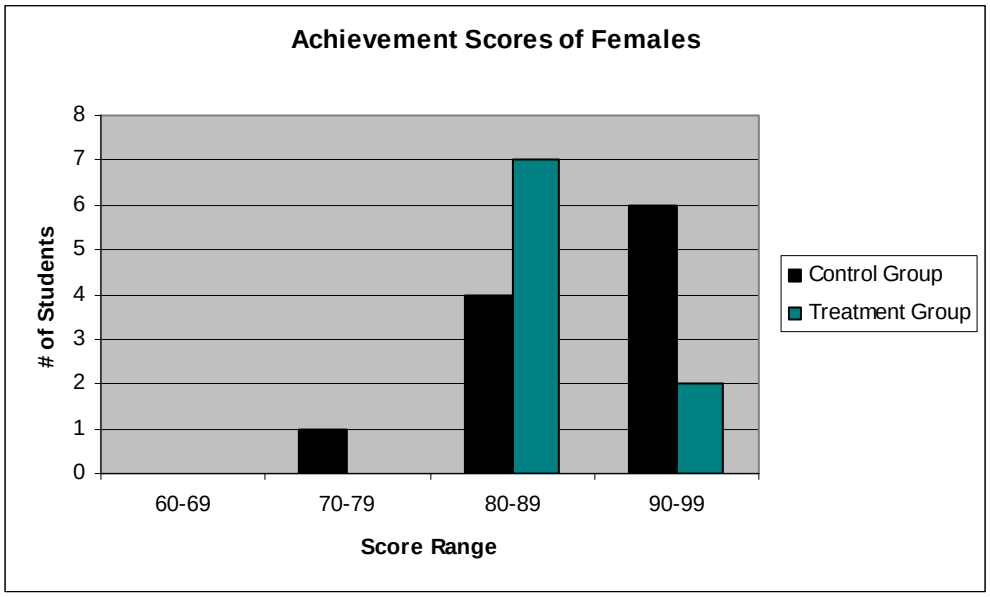

Figure 4.4

The researcher's second null hypothesis was that different methods of rubric implementation have no effect on achievement with regard to gender. A Student t-test at the .05 level of significance was conducted to test the hypothesis. First, the achievement scores of males in the control group and achievement scores of males in the treatment group were compared. The calculated value of the statistic for this test was -1.76. The critical value was 2.07. The degrees of freedom were 21. From this analysis, the researcher concluded that the data was

insufficient to reject the null hypothesis. In order to further analyze this hypothesis, the achievement scores of females in the control group and achievement scores of females in the treatment group were also compared. The calculated value of the statistic for this test was 0.53. The critical value was 2.10. The degrees of freedom were 18. From the analysis, the researcher again concluded that this data was also insufficient to reject the null hypothesis.

The graph in Figure 4.5 illustrates the essay achievement scores of the nineteen regular education students in the control group and the essay achievement scores of the nineteen regular education students in the treatment group. In the control group, one student received a percentage score of ninety-six, seven students received a ninety-two percent, three students received an eighty-eight percent, five students received an eighty-four percent, two students received an eighty percent, and one student received a seventy-six percent. The average score in the control group was eighty-eight percent. In the treatment group, one student also received a ninety-six percent, four received a ninety-two percent, seven students received an eighty-eight percent, six received an eighty-four percent, and one received an eighty percent. The average score of regular education students in the treatment group was eighty-nine percent.

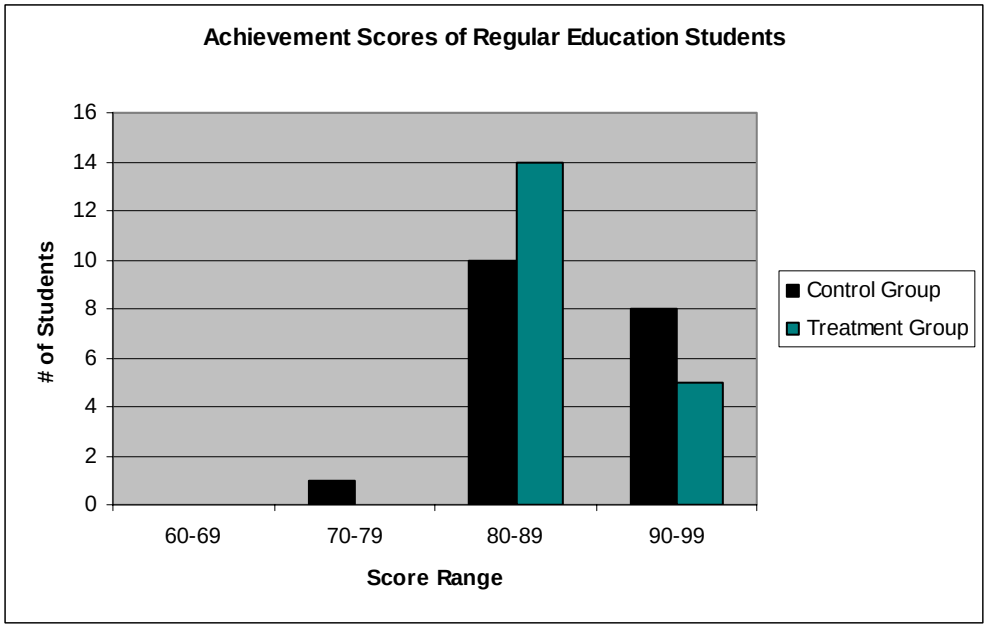

Figure 4.5

The researcher's third null hypothesis was that different methods of rubric implementation have no effect on achievement of regular education students. A Student t-test at the .05 level of significance was conducted to test the hypothesis. The achievement scores of

regular education students in the control group and achievement scores of regular education students in the treatment group were compared. The calculated value of the statistic for this test was -0.13. The critical value was 2.02. The degrees of freedom were 36. From this analysis, the researcher concluded that the data was insufficient to reject the null hypothesis.

The graph in Figure 4.6 illustrates the essay achievement scores of the two special education students in the control group and the three special education students in the treatment group. In the control group, one student received an eighty percent and the other received a seventy-six percent. The average score of special education students in this group was a seventy-eight percent. In the treatment group, one student received a ninety-six percent and two received an eighty-four percent. The average score of special education students in the treatment group was eighty-eight percent.

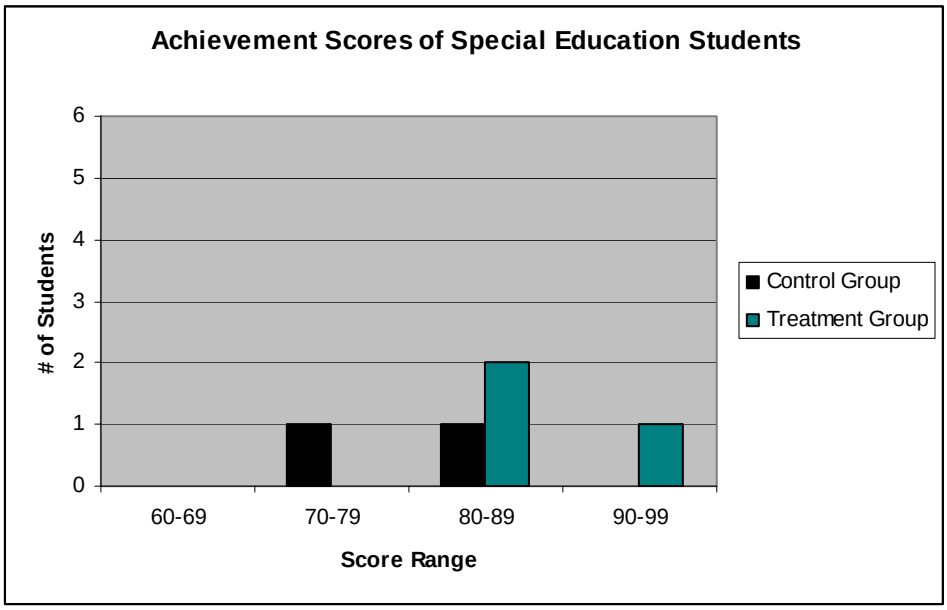

Figure 4.6

The researcher's fourth null hypothesis was that different methods of rubric implementation have no effect on achievement of special education students. A Student t-test at the .05 level of significance was conducted to test the hypothesis. The achievement scores of

special education students in the control group and the achievement scores of special education students in the treatment group were compared. The calculated value of the statistic for this test was -1.86. The critical value was 3.18. The degrees of freedom were 3. From this analysis, the researcher concluded that the data was insufficient to reject the null hypothesis.

Along with the essay achievement scores derived from the evaluation rubric, results were also analyzed from researcher designed pre and post surveys (Appendices A and C). Students read a series of statements regarding writing and assessment techniques and answered by circling either “4” (strongly agree), “3” (agree), “2” (disagree) or “1” (strongly disagree). Attitudinal scores were then determined by adding the six pre-survey statement totals together and adding the six post-survey statement totals together. Growth scores were calculated by subtracting the pre-survey total score from the post-survey total score.

The graph in Figure 4.7 illustrates the attitudinal growth of the twenty-one students in the control group and the attitudinal growth of the twenty-two students in the treatment group. In the control group, growth scores ranged from negative four to positive five. One student's growth score was five, one student's growth score was three, three students' growth scores were two, and three students' scores were one. Three students in the control group did not show any growth. In addition, six students had growth scores of negative one, one student had a growth score of negative two, two students had growth scores of negative three, and one student had a growth score of negative four. The average attitudinal growth score in the control group was -0.05. In the treatment group, growth scores varied from negative six to positive one. Six students' growth scores were one, three students did not show any growth, five students had scores of negative one, six students' scores were negative two, one student's score was a negative three, and one student had a growth score of negative six. The average attitudinal growth score for the treatment group was -0.9.

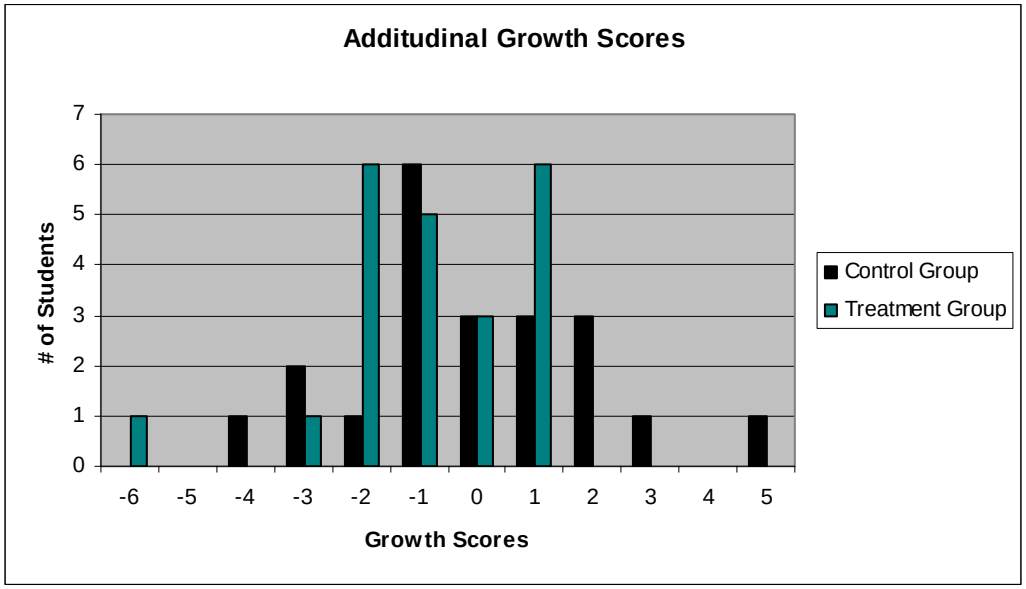

Figure 4.7

The researcher's fifth null hypothesis was that different methods of rubric implementation have no effect on student attitude. A Student t-test at the .05 level of significance was conducted to test the hypothesis. The attitudinal growth of the control group and the attitudinal growth of the treatment group were compared. The calculated value of the statistic for this test was 1.4. The critical value was 2.01. The degrees of freedom were 41. From this analysis, the researcher concluded that the data was insufficient to reject the null hypothesis.

The graph in Figure 4.8 illustrates the attitudinal growth of the ten males in the control group to the attitudinal growth of the thirteen males in the treatment group. In the control group, one male had a growth score of five, two males had a growth score of two, two had a growth score of one, three had growth scores of negative one, one had a growth score of negative two and one also had a growth score of negative three. The average attitudinal growth score of males in the control group was .3. In the treatment group, one male student had a growth score of five, two students had a growth score of zero, three students' growth scores were negative one, three students' growth scores were negative two, three students had a growth score of negative three, and one student's growth score was negative six. The average attitudinal growth score of males in the treatment group was -0.7.

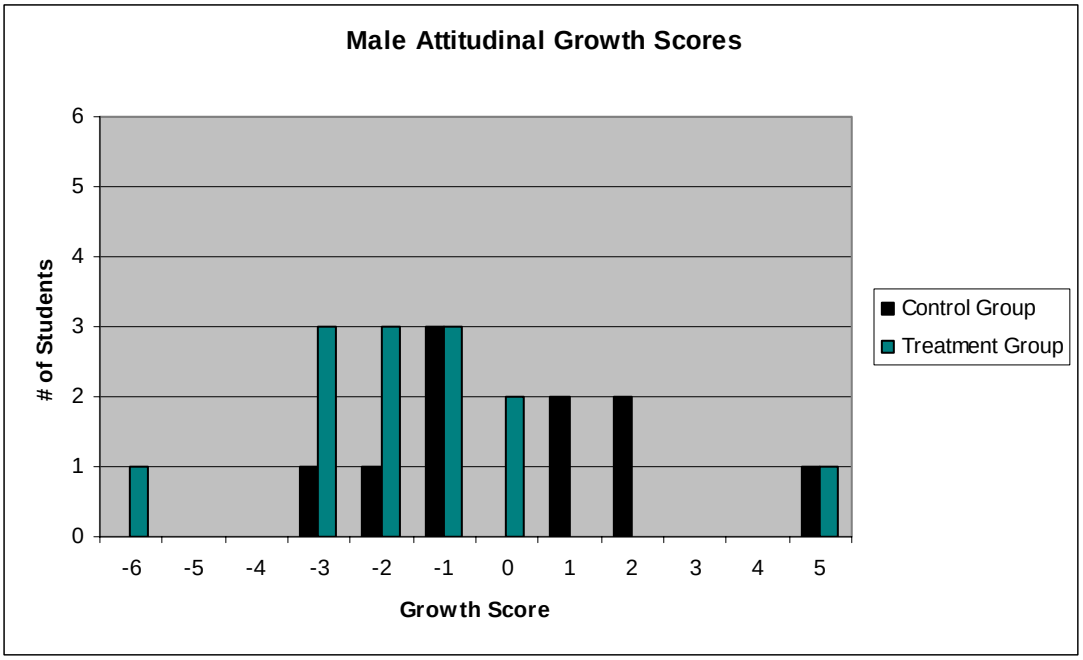

Figure 4.8

The graph in Figure 4.9 illustrates the attitudinal growth of the eleven females in the control group and the nine females in the treatment group. In the control group, one female had a growth score of three, one had a growth score of two, one had a growth score of one, three students didn't show any growth, three students' growth scores were negative one, one student's growth score was negative three, and one was negative four. The average attitudinal growth score for females in the control group was -0.3. In the treatment group, one female student had a growth score of one, one had a growth score of zero, three had a growth score of negative one, three had a growth score of negative two and one had a growth score of negative three. The average attitudinal growth score for females in the treatment group was -1.2.

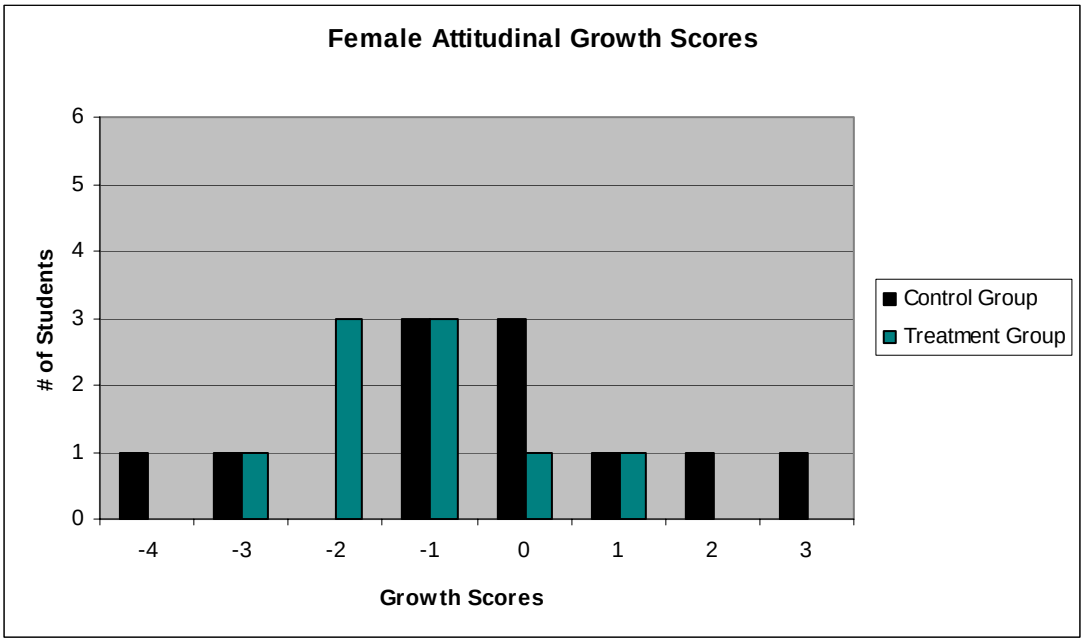

Figure 4.9

The researcher's sixth null hypothesis was that different methods of rubric implementation had no effect on attitude with regard to gender. A Student t-test at the .05 level of significance was conducted to test the hypothesis. The attitudinal growth scores of males in the control group and the attitudinal growth of males in the treatment group were compared. The

calculated value of the statistic for this test was 1.08. The critical value was 2.07. The degrees of freedom were 21. From this analysis, the researcher concluded that the data was insufficient to reject the null hypothesis. The attitudinal growth of females in the control group and the attitudinal growth of females in the treatment group were also compared to test this hypothesis. The calculated value of the statistic for this test was 1.1. The critical value was 2.10. The degrees of freedom were 18. From this analysis, the researcher concluded this data was also again insufficient to reject the null hypothesis.

The graph in Figure 4.10 illustrates the attitudinal growth of the nineteen regular education students in the control group and the attitudinal growth of the nineteen regular education students in the treatment group. In the control group, one student had a growth score of five, one had a growth score of three, two students had growth scores of two, two students had a growth score of one, three did not show any growth, six students showed growth scores of negative one, one showed a growth score of negative two, two had growth scores of negative three and one had a growth score of negative four. The average attitudinal growth score for the regular education students in the control group was -0.2. In the treatment group, four students showed growth scores of one, three did not show any growth, five showed scores of negative one, five also had scores of negative two, one student had a growth score of negative three, and one had a growth score of negative six. The average attitudinal growth score for the regular education students in the treatment group was -1.0.

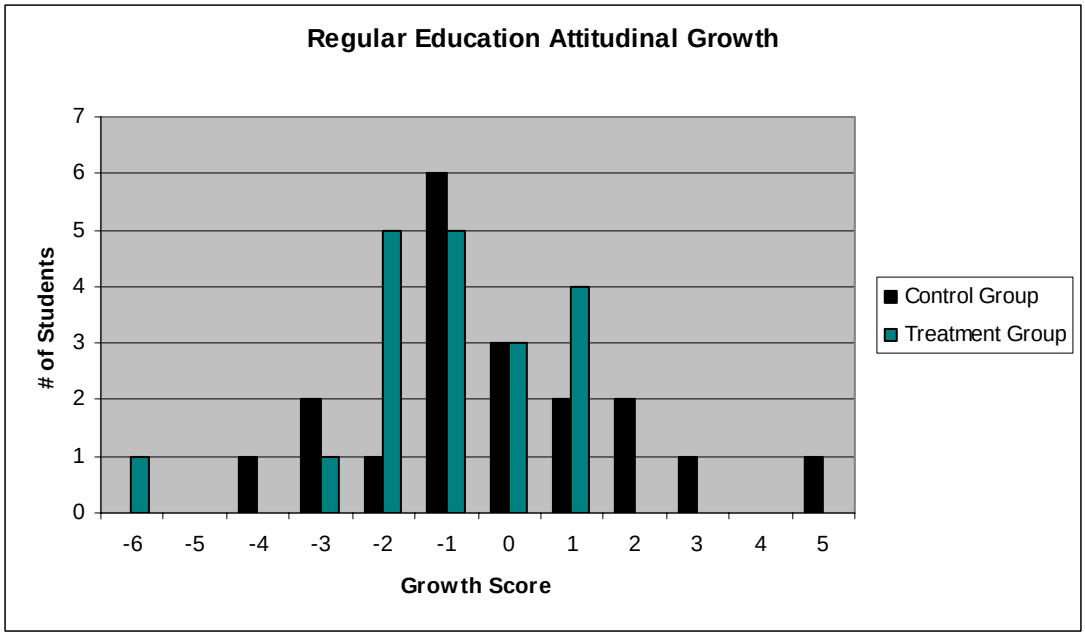

Figure 4.10

The researcher's seventh null hypothesis was that different methods of rubric implementation have no effect on the attitude of regular education students. A Student t-test at the 0.5 level of significance was conducted to test the hypothesis. The attitudinal growth of regular education students in the control group and the attitudinal growth of regular education students in the treatment group were compared. The calculated value of the statistic for this test was 1.3. The critical value was 2.02. The degrees of freedom were 36. From this analysis, the researcher concluded that the data was insufficient to reject the null hypothesis.

The graph in Figure 4.11 illustrates the attitudinal growth scores of the two special education students in the control group and the attitudinal growth of scores of the three special education students in the treatment group. In the control group, one student had a growth score of one and one student had a growth score of two. The average attitudinal growth score for the special education students in the control group was 1.5. In the treatment group, two students had a growth score of one and one student had a growth score of negative two. The average attitudinal growth score of the special education students in the treatment group was 0.

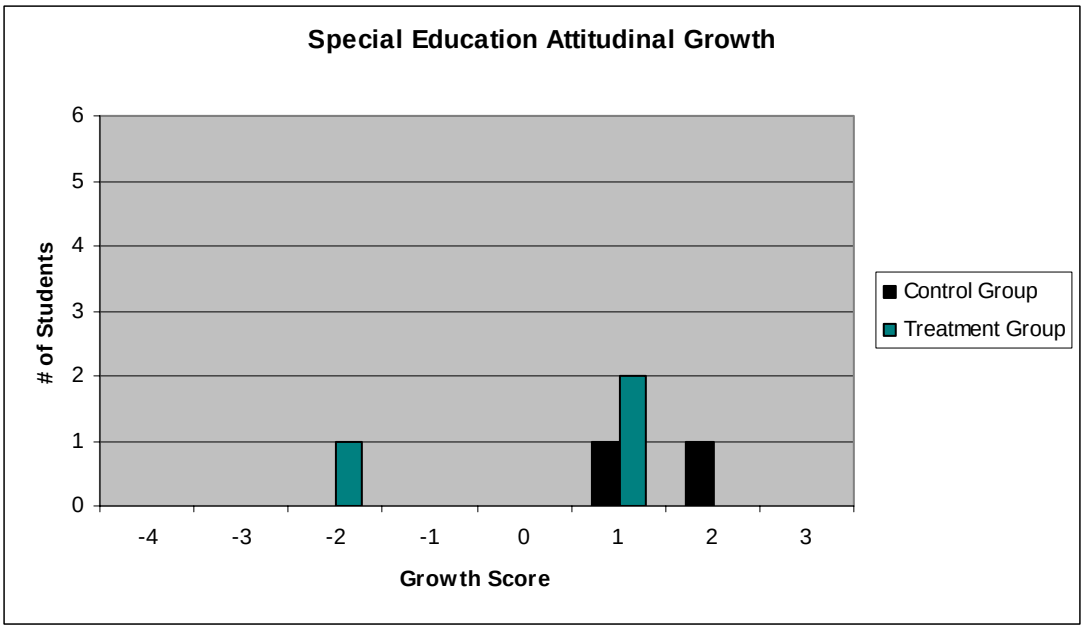

Figure 4.11

The researcher's final null hypothesis was that different methods of rubric implementation had no effect on attitude of special education students. A Student t-test at the .05 level of significance was conducted to test the hypothesis. The attitudinal growth of special education students in the control group and the attitudinal growth of special education students in the treatment group were compared. The calculated value of the statistic for this test was 1.1.

The critical value was 3.18. The degrees of freedom were 3. From this analysis, the researcher concluded that the data was insufficient to reject the null hypothesis.

Along with rating statements from strongly agree to strongly disagree, students in both the control group and treatment group answered four open-ended questions on the post-survey. The students were asked to respond to four questions dealing with rubrics as accurately as possible.

Students were asked in the first question to give their opinion of writing rubrics. The following nine responses were given from the twenty-one students in the control group. Four students responded that they liked rubrics because they tell how the assignment will be graded. Three students responded that a rubric helped them prepare and know what needed to go into a writing assignment. Two students responded that rubrics were a fair way of grading. One wrote that with a rubric, a writing assignment could be, "checked against the rubric if finished ahead of time." Three students responded that a rubric is an "okay" tool. One student responded that she liked rubrics, "because they tell the requirements," but she disliked them, "because they can limit writing as well." One student said a rubric was a very good grading process. Two students said rubrics helped them see how they did. Two students also mentioned rubrics were helpful, in general. Finally, two students indicated that rubrics told them what is required to earn an A grade.

The following eight responses were given from the twenty-two students in the treatment group for question number one. Four students responded that they liked rubrics because they showed what areas were done wrong. Two students indicated that a rubric shows them what they needed to improve on. Seven students said that rubrics helped them know what they needed to include in the writing assignment. One student said a rubric gave a fair grade. Three students

indicated it was any easy way to find out how they were graded and an easy way to understand their grade. Two students responded that they liked rubrics but did not give a specific reason. One student indicated that he liked rubrics because you get to see how you did in specific areas and are not just given a final grade. Finally, one student did not like rubrics because he felt they did not grade specifically. Overall, the students in both the control and treatment group seemed to have favorable opinions of rubrics.

Students were asked in the second question if the rubric helped them understand their grade on the tradition paragraph and why. The following responses were given from the control group. Twelve students indicated the evaluation rubric helped show what was wrong and what they needed to improve on next time. Four students thought it was helpful to see the rubric because it was clear and easy to see where the mistakes were. Three students responded that they did like the rubric, but did not give a reason why. One student responded that the rubric helped to point out spelling errors, and one student indicated that seeing the rubric was not helpful.

The following responses were given from the students in the treatment group for question number two on the post-survey. Thirteen students responded that seeing the rubric helped them determine what they did and did not have to improve on. Two students mentioned that the rubric showed them what they did well. Five indicated the rubric showed them what they did wrong. One student responded that the rubric was only somewhat helpful in understanding her grade, and one student indicated that he could use the rubric to try to get a better score next time. The majority of the students in both groups seemed to believe rubrics were helpful in understanding their grades on the essay.

Students were asked in question number three what some other uses for rubrics were. The following responses were given by the students in both the control and treatment groups. Sixteen students indicated rubrics would be helpful to use in science class on labs and experiments. More specifically, eleven students thought rubrics would be useful to grade science circuit boards, and three thought rubrics could be used for creating cell models. Eighteen students mentioned rubric use in reading class to assess book reports, fables or posters. Twenty-four students responded that rubrics could be used to grade news broadcasts, skits and research reports in English class. Many uses for rubrics, in various subject areas, were determined by the students.

Students were asked in question number four if they would like to be a part of developing rubrics for other assignments. The following answers were provided. In the control group, twelve students indicated affirmative responses. Six students responded negatively. One student responded "maybe". One responded that a whole day should be set aside for the class to make up a rubric for a writing assignment. One also indicated that she would like to "…put her ideas to work" by creating a rubric.

The following responses were provided by the treatment group regarding question number four. Nine students indicated they would like to be involved in developing rubrics for other assignments, and thirteen responded that they would not be interested in rubric development. Overall, about half of the sixth grade students in the research study would like to be a part of future rubric development.

Summary

The purpose of this study was to determine the impact of different methods of rubric implementation on achievement and student attitude. The quantitative data from the essay

achievement scores suggested that different methods of rubric implementation did not have an effect on achievement; consequently, the data was insufficient to reject the first null hypothesis. The students in the treatment group scored only slightly higher than the control group students as shown in Figures 4.1 and 4.2.

Different methods of rubric implementation also did not have an effect on achievement with regard to gender. Male students in the treatment group scored an average of two percent higher than the males in the control group on the tradition essay (Figure 4.3). The female students in the treatment group, however, scored an average of one percent lower than the female students in the control group on the essay (Figure 4.4). Both sets of data, as concluded by the researcher, were insufficient to reject the second null hypothesis.

In addition, different rubric implementation methods did not have an effect on achievement with regard to regular education students. Regular education students in the control group scored an average of one percent less than the regular education students in the treatment group on the tradition essay (Figure 4.5). Similarly, with regard to special education students, different rubric implementation methods again did not effect achievement. The average score of the special education students in the control group was seventy-eight percent, compared with eighty-eight percent in the treatment group (Figure 4.6). The researcher concluded that both sets of data again were insufficient to reject the third and fourth null hypotheses.

The data from the pre-survey and post-survey suggested that different methods of rubric implementation had no effect on student attitude. The average attitudinal growth score of the control group was -0.05, and the average attitudinal growth score of the treatment group was -0.9 (Figure 4.7). Therefore, the data, as concluded by the researcher, was insufficient to reject the fifth null hypothesis as well.

Different methods of rubric implementation also had no effect on attitude with regard to gender. In the control group, the average male attitudinal growth score was 0.3, and in the treatment group, the average male attitudinal growth score was -0.7 (Figure 4.8), both rejecting the sixth null hypothesis.

Furthermore, different methods of rubric implementation had no effect on attitude with regard to regular or special education students. The average attitudinal growth score of regular education students in the control group was -0.2. The average attitudinal growth score of regular education in the treatment group was -1.0. In addition, special education students in the control group had an average attitudinal growth score of 1.5, while the special education students in the treatment group had an average attitudinal growth score of zero. The researcher concluded that both sets of data were insufficient to reject the seventh and eighth hypotheses.

These results, accompanied by the qualitative data collected relating to rubric implementation and writing, indicated that there was not a significant impact on achievement and student attitude when different rubric methods are implemented. These findings were inconsistent with the major findings previously discussed on rubric implementation and writing achievement in chapter two. The rationale behind the findings in this study will be addressed in chapter five.

CHAPTER 5: CONCLUSION

Introduction

The purpose of this study was to assess the impact of different rubric implementation methods on achievement and student attitude. The design of the study was quasi-experimental, action research that produced both qualitative and quantitative data. Forty-three sixth-grade students wrote an essay on a favorite family tradition. An analytic evaluation rubric (Appendix B) was shown on an overhead transparency to the twenty-two students in the treatment group prior to drafting the essay. Each of the rubric's criteria and proficiency levels were also explained to the treatment group in detail. The twenty-one students in the control group did not view the evaluation rubric prior to drafting the essay; the rubric was only used to evaluate the essays of this group after final submission. Students in both groups were also given pre and post attitude surveys (Appendix A and Appendix C) designed by the researcher to measure student attitudes regarding writing and assessment techniques both before and after the treatment was given. The goal was to assess which of the two rubric implementation methods produced higher achievement levels and positive student attitudes. This chapter will discuss and interpret the findings from this study.

Discussion of the Findings

As indicated in chapter two, past research on rubric implementation suggested that there was a positive impact on writing achievement with the use of scoring rubrics in classrooms. Researchers have also found that achievement can increase when teachers share the evaluation rubrics with their students. Soles (2001) concluded that college students who knew the criteria set forth on the rubric ahead of time wrote more effectively than the students who did not. Andrade-Goodrich (1999) yielded similar results when she found that providing students with

instructional rubrics before constructing the first draft of an essay enabled them write significantly better. Both of these studies, however, contradicted the findings of this action research study as overall the different rubric implementation methods did not have an effect on achievement. The students in the treatment group who viewed the evaluation rubric scored only slightly higher on the tradition essay than the control group students who did not view it. This small increase in treatment group scores could have been because even though the students viewed the evaluation rubric on an overhead transparency, they did not have an actual copy to compare their initial drafts to once they were written. Similar achievement results occurred in regards to gender. Male students in the treatment group only scored an average of two percent higher than males in the control group. It is possible that this increase was because the male students in the treatment group knew the criteria set forth in the rubric and, consequently, were able to apply it to their drafts. Female students in the treatment group, however, scored slightly lower than females in the control group on the essay. The evaluation rubric may have been less motivating to the female students in the treatment group than the challenge of the grading unknown.

Other research cited in chapter two also found success implementing rubrics with regular education students. Ceprano and Garan (1998) determined a rubric was a useful method in charting conventions of language, Anderson and Woods (2002) found that students who used rubrics to self-evaluate became better performers, and Marzano (2000) found that science students scored higher on lab experiments with a rubric based assessment versus a traditional point based approach. The findings in this study contradicted these recent studies as the different rubric implementation methods did not have an affect on achievement of regular education students. Regular education students in the control group scored slightly less than the regular

education students in the treatment group on the tradition essay. This slight difference could have resulted because the treatment group students did not fully understand or remember the grading criteria and requirements on the rubric when drafting their essays. Some students may also have felt that they did not have a significant tradition to write about or elaborate on, affecting the overall scores. Students' exposure to rubrics, prior to this study, was also not prevalent.

In recent studies of special education students and rubric implementation, Superville (2001) learned that rubrics were helpful in increasing written expression of underachieving students, Schirmer et al. (1999) determined that using rubrics as a teaching strategy improved various writing areas for deaf students, and James and Abbott (2001) discovered that using a workshop approach to writing along with a six-trait rubric was effective in promoting writing growth of students with learning disabilities. These finding were also inconsistent with this study as different rubric implementation methods did not affect the achievement of special education students (Hypothesis #4). The average score of special education students in the control group was seventy-eight percent, compared with eighty-eight percent in the treatment group. The rubric viewing and explanation did seem to improve the achievement scores of the special education students in the treatment group, although not significantly. Another point to consider in analyzing the results of this group was that the number of special education students involved in the study was relatively small.

The data from the pre-survey and post-survey suggested that different methods of rubric implementation had no effect on student attitude. In fact, the overall average attitudinal growth score of the control group was -0.05, and the average attitudinal growth score of the treatment group was -0.9. Similar scores were collected on attitudinal growth of males (control = .3 and

treatment = -0.7), females (control = -0.3 and treatment = -1.2), regular education students (control = -0.2 and treatment = -1.0) and special education students (control = 1.5 and treatment = zero). The researcher was very surprised at the negative attitudinal growth results and believed these results were because on the post-survey, many students concentrated primarily on remembering what they had answered on the pre-survey, instead of just answering the questions honestly after the treatment period. Some students ended up guessing on the post-survey questions if they could not remember (which was not the objective of the post-survey) and growth scores may have been influenced.

On the open-ended survey questions, the students in both groups seemed to have favorable opinions of rubrics in general, and the majority of students thought that the evaluation rubrics were helpful in understanding their final essay grades. This correlated with Andrade-Goodrich's (1999) research who also determined from student questionnaires that students who had exposure to rubrics had a better understanding of how grades were given. Students in the study were also able to determine many uses for rubrics in a variety of subject areas. In addition, about half of the students indicated they would like to be a part of rubric development in the future.

The researcher believed that the strengths of this study were as follows: two relatively equal classes were compared, using a rubric to evaluate the essays minimized the objectivity of grading, students in both groups were given examples of completed tradition essays as well as a list of possible traditions to help generate ideas, and students in the treatment group were encouraged to use criteria on the rubric in the development of the essay.

One weakness of the study seemed to be the nature of the pre and post survey questions. The questions should have been specifically geared towards rubric implementation and traits of

quality writing. Different survey questions and more direction from the researcher in answering the questions possibly could have prevented students from answering “strongly agree” to a question on the pre-survey and “strongly disagree” on the same post survey question, for example. Another weakness was that students in the treatment group did not receive a hard copy of the evaluation rubric to aid in the development of the tradition essay. The researcher felt that this may have been a mistake on her part and negatively influenced achievement outcomes.

In addition to improving the weaker aspects of the study, if the researcher were to conduct this study again, several additional components would be changed. A third treatment group would be included in which the students had a role in the evaluation rubric’s development. Would students have higher achievement scores if they had more ownership in the evaluation? Secondly, as mentioned previously, the questions on the pre-survey and post-survey would be more specific and worded differently, hopefully resulting in positive growth scores instead of negative. Thirdly, English as a Second Language (ESL) learners would possibly be included as a separate group to study. Would these students benefit from seeing how a writing assignment was going to be evaluated ahead of time? Fourthly, the writing topic chosen for the study was not of interest to all of the students. Choosing a more popular topic may have had a more significant impact of achievement scores and student attitude in general. Also, although the students’ comments regarding rubrics were mostly positive, the researcher would have used rubrics more frequently before the study was conducted to better familiarize the students with various rubrics, their components, and how they could be used to achieve higher writing skills. Students, in addition, could have practiced grading sample essays with an evaluation rubric as well.

Summary

Contrary to the findings cited in chapter two, the findings in this study suggested that different rubric implementation methods did not have an impact on achievement in regards to gender, regular education students or special education students. The researcher found that achievement was not significantly affected when students in either of these groups viewed an evaluation rubric prior to writing an essay on a family tradition. Attitudinal growth findings resulting from the pre-survey and post-survey questions also suggested that different rubric implementation methods did not have an impact on overall student attitude as well.

Future studies involving different age levels and possibly another treatment group may be needed to determine whether or not different methods of rubric implementation are indeed effective in producing higher achievement scores and positive student attitudes. Using rubrics in younger grade levels or earlier in the year before a study is conducted may be necessary to familiarize students with possible criteria and proficiency levels in which their writing will be evaluated. It might be beneficial to repeat this study with a treatment group in which the students were engaged in the actual rubric development, providing students with an opportunity to develop assessment criteria, along with a control and treatment group similar to this study. This would be appropriate in determining whether that particular type of rubric implementation was more effective in increasing achievement scores and student attitudes.

Prior to conducting this study, the researcher strongly believed that showing a student an evaluation rubric prior to completing a writing assignment would allow the students to better evaluate their performance during the work phase and, in turn, produce a better quality of writing and higher achievement scores in general. It seemed apparent, however, that many students in the treatment group could not remember the criteria shown on the transparency in order to

compare their writing to the grading requirements set forth on the rubric. The researcher also believed that the majority of the students needed to have an actual copy of the evaluation rubric in hand to refer to while writing, as discussed previously.

To conclude, as stated at the beginning of this paper, rubrics have become increasingly popular across the nation as effective methods of evaluating students' performance. Whether the implementation method is developing a rubric with the students' input, showing a rubric prior to assignment completion to evaluate performance or as a way to award a final grade, rubrics can be an integral part of assessing student performance.

REFERENCES

Anderson, M., & Woods, D. (2002). Students designing and applying evaluation rubrics in an aerobics unit. Physical educator, 59 (1), 38-56.

Andrade-Goodrich, H. (1999). The role of instructional rubrics and self-assessment in learning to write: A smorgasbord of findings (Report No. TM-29). Montreal, Canada: American Educational Research Association. (ERIC Document Reproduction Service No. ED 431 029)

Ceprano, M., & Garan, E. (1998). Emerging voices in a university pen-pal project: Layers of discovery in action research. Reading research instruction, 38 (1), 31-56.

Gearhart, M., Herman, J., & Novak J. (1997). Establishing validity for performance based assessments: An illustration for collections of student writing. Journal of educational research, 89 (4), 220-233.

Jackson, C., & Larkin, M. (2002). Teaching students to use grading rubrics. Teaching exceptional children, 35 (1), 40-45.

James, Leigh Ann, & Abbott, Mary. (2001). How Adam became a writer: Winning strategies for low-achieving students. Teaching Exceptional Children, 33 (3), 30-41.

Johnson, R., Penny, J., & Johnson, C. (1998). A review of methods for resolving score discrepancies in the rating of performance assessments. Paper presented at the annual meeting of the South Carolina Educators for the Practical Use of Research, Columbia, SC.

Martin-Kniep, G. (1998). Standards, feedback and diversified assessment: Addressing equity issues at the classroom level. Reading and writing quarterly, 16 (3), 239-257.

Marzano, Robert. J. (2000). Transforming classroom grading. Alexandria, VA: Association for Supervision and Curriculum Development.

Penny, J., Johnson, R., & Gordon, B. (2000). Using rating augmentation to expand the scale of an analytic rubric. Journal of experimental education, 68 (3), 269-288.

Rhoads, Kyle. (1998). Kindergarten writing rubric (Report No. MF-01). Freeport, ME: Freeport Public Schools. (ERIC Document Reproduction Service No. ED 423 995)

Schafer, W. (2001). Effects of teacher knowledge of rubrics on student achievement in four content areas. Applied measurement in education, 14 (2), 151-171.

Schirmer, B., Bailey, J., & Fitzgerald, S. (1999). Using a writing assessment rubric for writing development of children who are deaf. Exceptional children, 65 (3), 383-397.

Simon, M., & Forgette-Giroux, R. (2001). A rubric for scoring postsecondary academic skills. Practical assessment, research and evaluation, 7 (18), 198-210.

Soles, Derek. (2001). Sharing scoring guides (Report No. PC-01). Denver, CO: Annual Meeting of the Conference on College Composition and Communication. (ERIC Document Reproduction Service No. ED 450 379)

Stuhlmann, J., Daniel, C., Dellinger, A., Kenton, R., & Powers, T. (1999). A generalizability study of the effects of training on teachers' abilities to rate children's writing using a rubric. Reading psychology, 20 (2), 107-127.

Superville, Linda. (2001). Oral assessment as a tool for enhancing students' written expression in social studies. Social studies, 92 (3), 121-126.

Wenzlaff, T., Fager, J., & Coleman, M. (1999). What is a rubric? Do practitioners and the literature agree? Contemporary education, 70 (4), 41-47.

Wiggins, G. (1998). Educative assessment: Designing assessments to inform and improve student performance. San Francisco, CA: Jossey-Bass.

Appendix A
Attitude Pre-Survey

Date _____

ID # _____

Writing Attitude Pre-Survey

Gender: Male Female

*** Please circle the number that best expresses your feeling about each of the following statements.**

1. I like to write when a teacher gives me a topic.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

2. I like to write when I get to choose the topic.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

3. I like to know how the writing assignment will be graded <u>before</u> I write it.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

4. I feel that being graded with a rubric is a fair way to evaluate writing.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

5. I think that seeing the rubric would help me better prepare my writing assignment.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

6. Reflecting on how I scored in each area of the rubric would be helpful for me to improve my writing skills in the future.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

*** Please respond to the following question.**

7. What is your opinion of a writing rubric?

Appendix B
Evaluation Rubric

Tradition Paragraph

Teacher's Name: Mrs. Smith

Student's Name: _____

CATEGORY	Excellent	Good	Satisfactory	Needs Improvement
Topic/ Closing Sentence	Topic sentence is clear, correctly placed, and is restated in the closing sentence.	Topic sentence is clear, correctly placed, and is not restated in the closing sentence.	Topic sentence is unclear or incorrectly placed, and is not restated in the closing sentence.	Topic sentence was not included in the paragraph.
Supporting Details	Paragraph has three or more supporting details that relate back to the main idea.	Paragraph has two supporting details that relate back to the main idea.	Paragraph has one supporting detail that relates back to the main idea.	Paragraph has no supporting details that relate back to the main idea.
Conventions	Paragraph has no errors in punctuation, capitalization, and/or spelling.	Paragraph has one or two punctuation, capitalization, and/or spelling errors.	Paragraph has three to five punctuation, capitalization and/or spelling errors.	Paragraph has six or more punctuation, capitalization and/or spelling errors.
Tradition Elements	Elements of the holiday tradition are expressed clearly and specifically.	Elements of the tradition are expressed.	Elements of the tradition are not expressed in detail.	Tradition elements are lacking.
Illustration	Student included a creative, detailed picture to reinforce the tradition.	Illustration somewhat reinforces the tradition.	Illustration is present, but not detailed.	No illustration was included.

Comments: **FINAL GRADE:**

Appendix C
Attitude Post-Survey

Date _____
ID # _____

Writing Attitude Post-Survey

Gender: Male Female

*** Please circle the number that best expresses your feeling about each of the following statements.**

1. I like to write when a teacher gives me a topic.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

2. I like to write when I get to choose the topic.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

3. I like to know how the writing assignment will be graded <u>before</u> I write it.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

4. I feel that being graded with a rubric is a fair way to evaluate writing.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

5. I think that seeing the rubric would help me better prepare my writing assignment.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

6. Reflecting on how I scored in each area of the rubric would be helpful for me to improve my writing skills in the future.

4	3	2	1
Strongly Agree	Agree	Disagree	Strongly Disagree

*** Please respond to the following questions as accurately as possible.**

7. What is your opinion of a writing rubric?

8. Did the rubric help you understand your grade on the tradition paragraph? Why?

9. What are some other uses for rubrics?

10. Would you like to be a part of developing rubrics for other assignments?

Appendix D
Research Data

Control Group - 6 Robinson

ID	Gender	R/Spec.Ed	Essay %	Pre #1	Pre #2	Pre #3	Pre #4
R 2	F	Reg	88	4	4	4	4
R 3	F	Reg	92	3	4	3	3
R 4	F	Reg	88	3	4	4	4
R 5	F	Reg	92	4	4	4	3
R 6	F	Reg	96	3	2	3	2
R 8	F	Reg	92	4	2	3	4
R 10	F	Reg	80	3	4	4	4
R 15	F	Reg	92	3	4	4	4
R 19	F	Reg	84	2	3	3	3
R 24	F	Reg	92	4	4	3	4
R 1	M	Reg	84	2	3	1	4
R 9	M	Reg	76	1	4	3	3
R 11	M	Reg	84	3	4	4	3
R 13	M	Reg	80	3	4	4	3
R 16	M	Reg	84	3	4	3	4
R 18	M	Reg	84	1	2	4	4
R 20	M	Reg	92	4	2	4	4
R 21	M	Reg	88	2	3	4	4
R 22	M	Reg	92	4	1	4	4
R 14	F	Special	76	4	1	3	3
R 17	M	Special	80	3	3	4	4

Treatment Group - 6 Smith

ID	Gender	R/Spec.Ed	Essay %	Pre #1	Pre #2	Pre #3	Pre #4
S 5	F	Reg	88	3	3	4	4
S 6	F	Reg	84	3	4	4	4
S 9	F	Reg	88	2	4	3	4
S 10	F	Reg	84	3	4	4	4
S 11	F	Reg	92	3	3	4	4
S 12	F	Reg	88	3	4	4	3
S 13	F	Reg	92	4	4	4	4
S 15	F	Reg	88	3	3	4	4
S 23	F	Reg	80	4	1	4	4
S 2	M	Reg	84	2	4	4	4
S 3	M	Reg	92	2	1	4	3
S 4	M	Reg	88	1	4	3	3
S 7	M	Reg	84	3	4	4	3
S 8	M	Reg	88	2	2	4	3
S 16	M	Reg	88	3	3	4	4
S 17	M	Reg	84	3	4	4	3
S 18	M	Reg	92	3	4	4	3
S 19	M	Reg	84	1	3	4	3
S 20	M	Reg	96	3	4	4	4
S 14	M	Special	84	2	3	3	2
S 22	M	Special	96	3	4	3	4
S 24	M	Special	84	3	4	3	4

Control Group - 6 Robinson

Pre #5	Pre #6	Total Pre	Post 1	Post 2	Post 3	Post 4	Post 5
4	4	24	4	4	4	4	4
4	4	21	3	4	4	4	4
4	4	23	3	4	4	3	4
4	4	23	3	3	4	4	4
4	3	17	3	2	4	3	4
4	4	21	4	2	3	3	3
4	4	23	3	4	4	4	4
4	4	23	3	3	4	3	3
4	3	18	2	3	3	3	3
3	4	22	3	4	3	4	4
2	3	15	2	4	4	3	4
2	3	16	1	4	3	3	3
4	3	21	3	4	4	3	3
4	4	22	3	3	3	3	3
4	4	22	3	3	4	4	3
4	4	19	2	3	4	4	4
4	4	22	3	2	4	4	4
3	4	20	2	4	4	3	3
4	3	20	4	1	4	4	4
4	4	19	4	2	4	3	3
3	3	20	3	4	4	4	4

Treatment Group - 6 Smith

Pre #5	Pre #6	Total Pre	Post 1	Post 2	Post 3	Post 4	Post 5
4	4	22	3	3	4	3	4
4	3	22	3	4	3	3	4
4	3	20	1	4	4	3	3
4	4	23	3	4	4	4	4
4	4	22	3	4	4	4	4
4	3	21	2	4	4	3	3
4	4	24	4	4	4	3	4
4	3	21	3	3	3	4	4
3	4	20	4	1	3	4	4
3	3	20	3	4	4	3	4
3	3	16	1	2	4	2	3
4	4	19	1	4	4	4	4
4	3	21	1	2	4	2	2
3	3	17	2	3	4	3	3
4	4	22	3	3	4	4	4
4	3	21	3	3	4	3	4
4	4	22	3	4	4	3	3
4	3	18	1	3	4	4	3
4	3	22	3	4	4	4	4
3	3	16	2	3	2	3	3
4	4	22	3	4	4	4	4
3	4	21	4	2	3	3	3

Control Group - 6 Robinson

Post 6	Total Post	Growth
4	24	0
4	23	2
4	22	-1
4	22	-1
4	20	3
3	18	-3
4	23	0
3	19	-4
3	17	-1
4	22	0
3	20	5
3	17	1
3	20	-1
4	19	-3
3	20	-2
4	21	2
4	21	-1
3	19	-1
4	21	1
4	20	1
3	22	2

Treatment Group - 6 Smith

Post 6	Total Post	Growth
3	20	-2
3	20	-2
3	18	-2
4	23	0
4	23	1
3	19	-3
4	23	-1
3	20	-1
3	19	-1
3	21	1
2	14	-2
4	21	-2
3	14	-6
2	17	0
3	21	-1
4	21	0
4	21	-1
4	19	1
4	23	1
4	17	1
4	23	1
4	19	-2