

SEX DISCRIMINATION IN THE EVALUATION
OF STUDENTS' WRITTEN COMPOSITIONS

by

NANCY F. KIEFER

A.B., Washington University, 1973

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Curriculum and Instruction

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1983

Approved by:

Mary McDonnell Harris
Major Professor

**THIS BOOK
CONTAINS
NUMEROUS PAGES
WITH THE ORIGINAL
PRINTING BEING
SKEWED
DIFFERENTLY FROM
THE TOP OF THE
PAGE TO THE
BOTTOM.**

**THIS IS AS RECEIVED
FROM THE
CUSTOMER.**

LD
2668
.TY
1983
K53
C.2

A11202 594250

ACKNOWLEDGMENTS

I dedicate this thesis to my niece, Stacey Renee Sternberg, for her help in writing the passages and with the hope that things will be different for her generation.

Research is not done in isolation. Several people helped in various aspects of the project. I thank my major professor, Dr. Mary Harris, and the other members of my committee, Dr. Nancy Smith, Dr. Robert Newhouse, and Dr. Harlan Trennepohl, for their help in obtaining subjects and their useful suggestions regarding the design of the experiment and the writing of the thesis. Appreciation is also expressed to Dr. Charles Rankin for his encouragement, support, and critical readings of the thesis throughout its development. Special thanks are due to my husband, Stephen, who provided continuous help from beginning to end and who has instilled within me a true appreciation for research.

TABLE OF CONTENTS

	<u>Page</u>
LIST OF FIGURES	iii
LIST OF TABLES	iv
CHAPTER	
I. INTRODUCTION	1
Purposes and Objectives	2
Significance of the Study	2
Definition of Terms	3
Delimitations of the Study	5
II. REVIEW OF THE LITERATURE	6
Summary	15
III. RESEARCH METHODOLOGY	18
Developmental Stages	18
Pilot Study 1	18
Pilot Study 2	20
Experiment	22
Subjects	22
Materials	22
Procedure	25
Data Analysis	25
IV. RESULTS	28
V. CONCLUSIONS AND RECOMMENDATIONS	31
Conclusions	31
Discussion	32
Recommendations	34
REFERENCES	37
APPENDICES	41
A. Passages Evaluated by Subjects in Pilot Study 2 and Experiment	42
B. Rating Sheet for Pilot Study 2	46
C. Cover Page from Materials Given to Subjects	48
D. Rating Sheet Used in Experiment	50

LIST OF FIGURES

Figures	<u>Page</u>
1. Mean Ratings of "My Saddest Day" By Sex of Author	27
2. Mean Ratings of "My Car Trip" by Sex of Author	28
3. Mean Ratings of "Junior High" by Sex of Author	30

LIST OF TABLES

Table	<u>Page</u>
1. Passage Values	19
2. Name/Passage Combinations	24

Chapter 1

INTRODUCTION

Sex bias in the classroom has been a topic of growing concern among educators in recent years. Much of the research relating to this topic has dealt with verbal interactions between teachers and students. For example, Hall and Sandler (1982) found that from preschool to college, male students were shown to have more interactions with teachers than female students. Not only did female students receive less academic contact, but they were asked fewer high level questions and given less positive feedback for correct answers than males (Sadker & Sadker, 1982). On the other hand, several studies have revealed definite differential treatment given to students which has been damaging to males. Males were more likely to be reprimanded in classrooms, even when the observed behavior of males and females did not differ. Also, males were more likely to be referred to school authorities for misconduct than were females (Sadker & Sadker, 1982). In each of the cited examples, discrimination based on gender alone was found to exist in a particular aspect of the educational setting.

Purposes and Objectives

The purpose of the present study was to examine for sex discrimination another aspect of school life, the evaluation of student academic performance. Specifically, the objective was to determine if teachers were sex biased when evaluating the written work of elementary school students. Further, were teachers' evaluations influenced by an interaction between the gender of the student and the sex role behavior conveyed by the student's written composition?

To test this, elementary and secondary teachers were given three compositions purportedly written by sixth grade students to evaluate on a number of dimensions. Each passage was headed by either a male name or a female name and conveyed one of three different sex role behaviors. The teacher ratings of the passages were compared to determine if a difference existed between the evaluations of a passage thought to be written by a male and the evaluations of the same passage thought to be written by a female. Any differences in the ratings would reflect sex bias as the works were identical except for the gender conveyed by the author's name on the composition.

Significance of the Study

Although much research has been conducted in the area of sex bias in evaluation of performance in non-academic settings, few studies have examined this issue in the academic realm.

Two studies that have revealed sex bias in evaluation have been conducted in Australia at the high school level (Bernard, 1979; Bernard, Elsworth, Keefauver, & Naylor, 1981). Bernard found that individuals exhibiting male sex role behaviors were given higher evaluations and thought to have more potential for academic success. Although this study revealed significant sex bias at the high school level, these findings cannot be automatically generalized to the American school context at the elementary level.

Further, research has already identified sex discrimination in classroom interactions and differential treatment given to male and female students (e.g. Sadker & Sadker, 1982). The present study examined whether sex bias in the evaluation of performance existed at the elementary school level. It was important to determine if, in fact, boys and girls were evaluated differently especially when they were demonstrating role behavior that was considered atypical for their gender. By elucidating the situations where sex bias exists and incorporating them within a framework of previous work, one can then initiate the process for change.

Definition of Terms

Sex bias is differential treatment given to individuals based on their gender alone. In the present study, sex bias is defined operationally as differential ratings given to male and female students on identical compositions.

Written performance is defined as a one paragraph passage on a particular topic conveying a specific sex role behavior.

Male sex role behavior. Those actions perceived as being masculine according to one's stereotypes of appropriate behavior for that gender.

Female sex role behavior. Those actions perceived as being feminine according to one's stereotypes of appropriate behavior for that gender.

Neutral sex role behavior. Those actions which can be perceived as being neither masculine nor feminine according to one's stereotypes of appropriate behavior for both genders.

Stereotype. A stereotype is a standard, oversimplified opinion held in common by many members of one group about another (Boyer & Boyer, 1981).

Control group refers to those subjects who received passages to evaluate without identification of the student author.

Experimental group refers to those subjects who received passages to evaluate which included the student author's name, a name that clearly conveyed the gender of the student.

Evaluation criteria. Those dimensions on which subjects rated the passages, specifically: creativity, neatness, effort, presentation of ideas, and overall evaluation.

Delimitations of the Study

The data in the present study reflect the evaluation of sixth grade work by elementary and secondary teachers in one Midwestern university. Generalizations concerning sex bias in evaluation of performance at other grade levels and in other locations must be made cautiously. Clearly there is a need for further research on sex bias in the classroom. This study provides information concerning this issue at one level.

Chapter 2

REVIEW OF THE LITERATURE

Evaluation of performance is a ubiquitous part of a person's life. From early in life through adulthood, one's performance is judged by parents, peers, teachers, employers, and significant others. Recent research has shown that the identical performances of males and females are not always evaluated in the same manner. Prejudicial evaluation based on gender alone can be damaging to both males and females.

The classic study showing an evaluative bias against females was conducted by Goldberg in 1968. Female subjects were given what they thought were published journal articles from 6 professional fields; 2 traditionally masculine, 2 traditionally feminine, and 2 neutral fields. The subjects rated these articles on writing style, professional competence, professional status, and ability to sway the reader. For each article, half of the subjects saw John T. McKay as the author and half saw Joan T. McKay as the author. On all criteria, the articles thought to be authored by John T. McKay were rated more favorably than identical articles thought to be authored by Joan T. McKay. This was true in all professional fields, even the ones considered traditionally feminine.

The results of this landmark research inspired several subsequent studies that employed the same paradigm of evaluating a work by a fictitious person. Pheterson, Kiesler, and Goldberg (1971) designed a study to investigate the conditions under which women were prejudiced against women. Subjects judged paintings on the measures of competence and predicted future success of the artist. Half of the subjects thought that the artist of the painting was a female, half thought the artist was a male. In addition, half of the subjects thought the painting was a contest entry, the other half thought it was a winner. Interestingly, when the paintings were thought to be entries in a contest, female works were evaluated less favorably than the male's identical work. However, winning paintings were not judged differently based on gender of the assumed artist. The authors concluded that women who are attempting to accomplish are judged less favorably than men, whereas women who have achieved success are evaluated the same as men.

Sex bias has also been shown to be a function of the perceived competence level of the individual being evaluated. In a study by Deaux and Taynor (1973), college students rated applicants for a study abroad scholarship program. Male and female subjects listened to taped interviews and rated the applicants on competence and intelligence. Results showed that highly competent males were rated more positively than highly competent females. However, males of low competence

were rated lower than similar females. The gender of the rater was not significant in the evaluation of the applicants. Thus, it appears a complex interaction exists between competence level and gender of the person being evaluated.

A similar finding was reported by Heneman (1977) who had college students evaluate hypothetical applicants for the job of life insurance agent. High scoring females were rated as being less suitable for the job than high scoring males. Again, at the other end of the scale, evaluators were more confident about hiring low-scoring females than low-scoring males.

In yet another study (Lao, Upchurch, Corwin, & Grossnickle, 1975), a particular bias was found when the females were acting in "inappropriate" sex roles. Male and female college students viewed videotapes of males and females in high, medium, and low assertive roles. Subjects rated the role players on intelligence and likeability. Males were judged to be more intelligent and likeable than females especially in highly assertive roles.

Although the studies cited above differed somewhat in their emphasis, they all showed that when evaluating performance, both males and females were biased against women. Several variables such as competence and appropriateness of sex role were found to influence this bias. In contrast to these findings, other studies have shown that in certain situations females were rated more favorably than males given equal performances.

In one study (Taynor & Deaux, 1973) male and female subjects read descriptions of males and females performing well in an emergency situation. It was found that females were rated as more deserving of reward than males for equivalent performance. A similar sex bias occurred when males and females read a description of either a male or female actor performing a task (Taynor & Deaux, 1975). Women performing a masculine task were rated as more deserving of reward than equally performing males. Males were not overevaluated in a similar fashion on the feminine task.

The effect of one's competence level on evaluation bias was examined by Bigoness (1976). Potential employers rated performance standards of males and females. Results indicated that low-performing males and low-performing females were rated about the same. However, high-performing females were rated significantly higher than high-performing males.

The results of the Taynor and Deaux (1973; 1975) and Bigoness (1976) studies were in direct contrast to those which reported lower evaluations of competent females. One possible reason for this discrepancy is explained by the equity theory (Adams, 1965; Leventhal & Michaels, 1971). Briefly, the equity theory states that the amount of rewards received is determined by the circumstances of the input to the situation. When non-voluntary constraints (something a person cannot control: sex, age, height, etc.) limit the input, the rewards are adjusted accordingly to achieve a sense of balance or equity. In the

Taynor and Deaux (1973) experiment, then, being a woman in a masculine situation was considered to be a nonvoluntary constraint, and accordingly, a woman was rewarded more than a man for identical performance. Similarly, in the Bigoness study, women were perceived as performing well in unexpected contexts and thus were overevaluated for their identical performance.

To summarize, in several different nonacademic situations such as evaluating art, job applications, performance in an emergency situation, and role playing, sexual bias was shown to be a significant factor in the evaluation of identical performance. As one might expect, the differential treatment given to males and females does not begin and end in the non-academic world. It actually has been documented to occur early in one's schooling and continues through college.

From preschool to college, male students were shown to have more interactions with teachers than female students (Hall & Sandler, 1982). Cherry (1975) found that female preschool teachers verbally interacted more, verbally initiated more, and used more attentional-marked utterances with boys than with girls. Attentional-marked utterances are statements which begin with words such as hey, see, now, no, OK, or listener's name; words which are thought to catch or hold the listener's attention.

Dweck, Davidson, Nelson, and Enna (1978) observed fifth graders and found that boys received more negative feedback than girls. This feedback was received for failure to follow

the rules of form or for misconduct. The positive feedback that boys received was for intellectual competence. In contrast, girls received positive feedback for intellectually irrelevant aspects of work such as effort or neatness and negative feedback for intellectual performance. These observations formed the basis for an experimental study which tested the effects of feedback on self perception of ability.

Dweck et al. (1978) hypothesized that failure feedback based only on correctness of answers would be seen by students as indicative of a lack of ability. Conversely, students given failure feedback that was solution irrelevant would be more likely to view failure as a reflection of their effort or the evaluator. In their experiment, students were given word puzzle problems to solve. Upon completion of the task, children were either given failure feedback based only on the correctness of the solution or failure feedback which was solution irrelevant. These types of feedback simulated teacher-girl and teacher-boy conditions, respectively, as observed in their first study. Students then were asked to select one of three reasons for not doing well on the puzzles: 1) I did not try hard enough. 2) The experimenter was too fussy. 3) I am not very good at it. Analysis of the data revealed that regardless of sex, children who received failure feedback that was solution specific were much more likely to view failure as a lack of ability. Students receiving solution irrelevant failure feedback were more likely to attribute that failure to a lack of effort or the fault of the evaluator.

Dweck et al. concluded that the differential interactions between teachers and students influenced the pupils' perceptions of why they failed. Males attributed failure to lack of effort whereas females attributed failure to a lack of ability. Extending these data to a success situation, one would predict that a male's success results from ability and a female's success results from luck or effort. This hypothesis was confirmed, in fact, by Deaux and Emswiller (1974) who showed that equivalent performances by males and females were attributed to different causes. Whether the required task was masculine or feminine (identifying tools or household objects, respectively) subjects attributed a male's success to skill and a female's success to luck or effort. These results provide further support for the equity theory in explaining sex bias in the evaluation of identical performance.

Differential treatment of boys and girls in school was also noted in seventh and eighth grade where boys were found to interact more frequently with teachers (Good, Sikes, & Brophy, 1973). Again, a given contact was more likely to be negative for boys than for girls. At the college level, males also engaged in proportionately more student-teacher interactions than female students in male-taught classes. However, there was no sex difference in female-taught classes (Sternglanz & Lyberger-Ficek, 1977).

Certain studies found, too, that differential treatment involved an interaction between gender of the student and other

independent variables. For instance, Good et al. (1973) showed that high-achieving boys exceeded other boys and all girls in the frequency and quality of teacher contacts. Low-achieving boys received the poorest contacts with both male and female teachers. One is reminded of the Deaux and Taynor (1973) study and Heneman's (1977) findings in which highly competent males were judged more favorably than highly competent females, and males of low competence were rated lower than similar females.

The verbal interaction that occurs between teachers and students is just one way in which sex bias is apparent in the classroom. Studies have shown that students who have certain personality characteristics are more preferred by teachers. Feshbach (1969) asked student teachers to rate story situations depicting boys and girls with various personality clusters. On intellectual and social dimensions, student teachers rated significantly more positive the conforming, rigid, dependent child as compared to the flexible, nonconforming, independent child. In fact, student teachers perceived most positively the rigid, conforming girl and secondly, the rigid, conforming boy. The lowest ratings were given to the independent, assertive girl (see also Etaugh & Hughes, 1975). Recall the Lao et al. (1975) study which showed that the lowest ratings in likeability and intelligence were given to females in highly assertive roles.

The preference for certain personality characteristics probably affects a teacher's evaluation of student performance in the classroom. McCandless, Roberts, and Starnes (1972)

found that elementary teachers gave girls higher grades than boys even when no sex differences were observed in their standardized achievement scores. Further, according to Sadker and Sadker (1982), girls are less likely than boys to be identified as having learning and reading disabilities, and to repeat grades in school. In these cases, preference for rigid, conforming female behavior may have been a factor in the teacher evaluation of student performance.

Bernard (1979) examined the tendency to give females higher evaluations at the high school level and found quite different results. He had teachers rate descriptions of male and female high school students who demonstrated either masculine or feminine sex role behavior. Data showed that students with masculine sex-role behavior were rated as more competent than ones with feminine sex-role behavior. In addition, an essay thought to have been written by a student with masculine sex-role behavior was rated higher in quality than the same essay thought to have been written by a student with feminine sex-role behavior.

In a follow-up study (Bernard et al., 1981), teachers read descriptions of a male or female high school student in bidimensional sex roles (high or low masculinity paired with high or low femininity). The teachers were then asked to evaluate the sex-role description of the student on a number of dimensions. Again, data indicated that high levels of masculinity rather than femininity, independent of the gender

of the student, were associated with academic success and university potential.

After evaluating the sex-role description, the same teachers read and rated an essay thought to be written by Ian or Mary. On 5 of 16 dimensions (individuality, freshness, defend beliefs, express point of view, and self-confidence), the essay thought to be written by a highly masculine student was rated more positively than the one thought to be written by a student with low masculinity. However, on the remaining dimensions, no sex bias was observed in the ratings of the essay. Similar findings were reported by Duval (1980) in which no differential grading behavior was found for male and female mathematics students in high school. Bernard et al. (1981) suggest that the similarity of ratings indicates that the observed higher grading of female students by elementary school teachers may disappear at the high school level.

Summary

Differential treatment of males and females has been shown, then, to occur in many academic and nonacademic settings. Whether subjects were rating research articles, works of art, scholarship or employment applications, or videotapes of performance, bias based on gender alone was documented. In certain cases, as discussed earlier, bias was exhibited as a means of equalizing one's perceptions of what should occur with what was actually observed. Thus, in the Taynor and Deaux

(1973) and Bigoness (1976) studies high performing females were overrated on masculine perceived tasks.

As mentioned earlier, there are interesting parallels that can be made between what happens in the academic and nonacademic realms of experience. The finding of Dweck et al. (1978) concerning perceived reasons for failure were mirrored by Deaux and Emswiller's (1974) findings of perceived reasons for success for males and females. Specifically, a male's success is due to skill whereas a female's success may be attributed to luck or effort. In addition, the Good et al. (1973) results showed that high-achieving boys received the best contacts with teachers whereas low-achieving boys received the poorest contacts. These findings were paralleled in nonacademic situations according to the results of Deaux and Taynor (1973) and Heneman (1977) which showed that highly competent males were judged more favorably than highly competent females, and males of low competence were rated lower than similar females.

Finally, Feshbach's (1969) study of personality characteristics most preferred by teachers showed the lowest ratings given to the independent, assertive female. These findings were similar to those in the Lao et al. (1975) study which showed that females in highly assertive roles were given the lowest ratings in likeability and intelligence.

Clearly the messages one receives in the classroom are powerful. The internalization of the perceptions of self and

of others learned in school carries over to adulthood when one is put in the position of evaluator. In that role, a person makes judgments and attributions based on these learned perceptions, thus perpetuating the cycle.

Based on the recurring patterns of bias, further study in the specific area of sex bias in evaluation of performance was recognized. Bernard (1979) and Bernard et al. (1981) studied this issue at the high school level in Australia. Their data indicated that high levels of masculinity rather than femininity were valued and rewarded more positively. Using a similar paradigm, the present study examined sex bias in evaluation at the elementary school level.

Chapter 3

RESEARCH METHODOLOGY

The purpose of this study was to determine if teachers were sex biased when evaluating the written work of elementary school students. Three compositions (see Appendix A), each conveying different sex role behaviors, were evaluated on a number of dimensions by elementary and secondary teachers.

Developmental Stages

Prior to the actual data collection for the present experiment, two developmental stages were taken. The first phase involved the selection of paragraphs to be used as the samples of written work. These passages were constructed to be equal in length, T-unit number, T-unit length, and ideation fluency (see Table 1). The second phase consisted of two pilot studies.

Pilot Study 1

The first pilot study was conducted to identify students whose handwriting was perceived as being neutral (neither feminine nor masculine) in style. Students having the most neutral handwriting were selected to write the passages for the actual study. This was deemed necessary so that the gender

Table 1
Passage Values

Passage	Total number of words	Number of sentences	Number of T-units	T-unit length
My Saddest Day	64	4	6	10.6
My Car Trip	62	5	6	10.4
Junior High	61	4	6	10.1

of the student indicated by the name of the author was not contradicted by the style of handwriting.

Subjects. Forty-eight male and female subjects were selected from upper division education classes at Kansas State University.

Materials/Procedures. Subjects received xeroxed copies of a paragraph written by sixteen different sixth grade students from Salina, Kansas. Subjects also received a rating sheet containing 16 scales, each ranging from 1 through 9. The number 1 was identified as representing very feminine handwriting, the number 5 represented neutral or difficult to classify, and the number 9 represented very masculine handwriting. Subjects were given 15 minutes to rate each handwriting by circling the number on the scale that best reflected their opinion.

Data Analysis. Means were computed for each handwriting sample. Those individuals whose average was closest to neutral (5) were selected to write the paragraphs.

Results. The means for the 16 handwriting samples ranged from 2.3 to 6.7. The three students selected had mean scores of 5.2, 5.4, and 5.6.

Pilot Study 2

The second pilot study examined the passages for the sex role behavior which they conveyed. For the purpose of the study it was important that one passage convey masculine sex

role behavior, another convey neutral sex role behavior, and a third convey feminine sex role behavior.

Subjects. Fifteen male and female subjects were selected from upper division education classes at Kansas State University.

Materials/Procedures. Subjects received xeroxed copies of three compositions written by sixth grade students (see Appendix A). The order of passage presentation was systematically varied for the subjects. One passage described an emotional reaction to the death of a grandmother (My Saddest Day), another described selecting classes for junior high school (Junior High), and a third described a car trip (My Car Trip). The authors' names were not written on the passages.

Subjects also received three rating sheets. Each included a designated place for the title of the passage and five rating scales (1-9) for the following criteria: creativity, neatness, presentation of ideas, effort, and overall evaluation. An additional scale (1-9) was included for subjects to indicate the sex role behavior conveyed by each passage (see Appendix B). Subjects were instructed to circle the number on the scale that best represented how they felt about the passage on each dimension.

Data Analysis. Statistical analysis was performed only for the last scale which reflected the sex role behavior conveyed by the passage. Analysis of variance and appropriate post hoc tests were used to determine if the three passages were significantly different.

Results. The mean ratings for each passage along the scale which indicated the sex role behavior conveyed were as follows: My Saddest Day, 6.0; My Car Trip, 4.9; and Junior High, 3.1. Statistical analysis of the ratings indicated significant differences between the passages, $F(2,28) = 22.8$, $p < .001$. Post hoc comparisons using the Neuman-Keuls procedure revealed that the passage My Saddest Day was rated significantly higher (toward the feminine end of the scale) than the neutral passage, My Car Trip ($p < .05$). In addition, the passage Junior High was rated significantly lower (toward the masculine end of the scale) than the neutral passage ($p < .05$). It was concluded that each passage did convey a different sex role behavior (feminine, neutral, masculine) as required by the design of the experiment.

Experiment

Subjects

One hundred and twenty male and female subjects were randomly selected from graduate education classes at Kansas State University. A minimum of one year teaching experience was necessary to be included in the study.

Materials

Each subject received a packet of materials containing a cover page, three passages, and three rating sheets.

The cover page included personal data questions to be answered by each subject and the following explanation stating

the purpose of the study: The purpose of the present research is to determine how teachers evaluate the quality of written work done by sixth grade students (see Appendix C).

Inside the packet, the subject found xeroxed copies of handwritten compositions done by 6th grade students (see Appendix A). As indicated by pilot data, one passage conveyed female sex role behavior (My Saddest Day), another conveyed neutral sex role behavior (My Car Trip), and a third conveyed male sex role behavior (Junior High). Again the order of the passages within the packet was systematically varied.

For experimental subjects, each of the three passages had either a male name or a female name on the composition. A summary of possible name/passage combinations is shown in Table 2. Control subjects received the same passages with no student name.

Three rating sheets were included in the packet of materials. For the experimental subjects, a place was designated for filling in the title of the passage and the student author (see Appendix D). Having the subjects write down the name of the student author assured awareness of the child's gender. Control subjects' rating sheets provided only a space for the title of the passage. Subjects then rated the passages on five dimensions: creativity, neatness, effort, presentation of ideas, and overall evaluation by circling the appropriate number (1-9) on each scale.

Table 2
Name/Passage Combinations

Passage	Male name	Female name	Control
My Saddest Day	Mike	Julie	No name
My Car Trip	Bill	Laura	No name
Junior High	Gary	Pam	No name

Procedure

The experimenter administered the task in education classes at Kansas State University. Each class was designated as either an experimental group or a control group. For the control group (n=40), all subjects received identical passages with no student names. For the experimental subjects (n=80), the eight possible name/passage combinations were randomly distributed to the subjects. A brief introduction was made before distributing the materials. Subjects then had 20 minutes to read and rate the passages before the materials were collected by the experimenter.

Data Analysis

Each passage/name combination was rated 40 times. Analysis of variance was used to determine if a difference existed between the evaluations of passages thought to be written by a male and the evaluation of passages thought to be written by a female. Control passages (those with no name) provided a baseline evaluation. In addition, the binomial test (Siegel, 1956) was used to determine if a gender difference existed in ratings across passages.

Chapter 4

RESULTS

The mean scores for the passage *My Saddest Day*, which conveyed feminine sex role behavior, are shown in Figure 1. Analysis of the data indicated a significant author effect, $F(2,117) = 5.34$, $p < .01$. There was also a significant evaluation criteria effect, $F(4,468) = 47.17$, $p < .001$, but no significant author by evaluation criteria interaction, $F(8,468) = .43$, $p > .10$. As can be seen from Figure 1, subjects evaluating the passage thought to be authored by Mike rated it consistently higher than those evaluating the identical passage with no author name. Conversely, the passage purported to be written by Julie was rated consistently lower than the same passage with no author name (control).

Figure 2 presents the mean scores for the passage *My Car Trip* which was shown by pilot data to convey neutral sex role behaviors. The only significant effect was that of evaluation criteria, $F(4,468) = 25.80$, $p < .001$. The author effect approached significance, $F(2,117) = 2.36$, $p = .09$, while the interaction between author and evaluation criteria was nonsignificant, $F(8,468) = 1.03$, $p > .10$. A relatively consistent pattern observed in the data was that Laura received lower ratings than Bill on four of the five criteria. The

**THIS BOOK
CONTAINS
NUMEROUS PAGES
WITH DIAGRAMS
THAT ARE CROOKED
COMPARED TO THE
REST OF THE
INFORMATION ON
THE PAGE.**

**THIS IS AS
RECEIVED FROM
CUSTOMER.**

Figure 1
Mean ratings of "My Saddest Day"
by sex of author

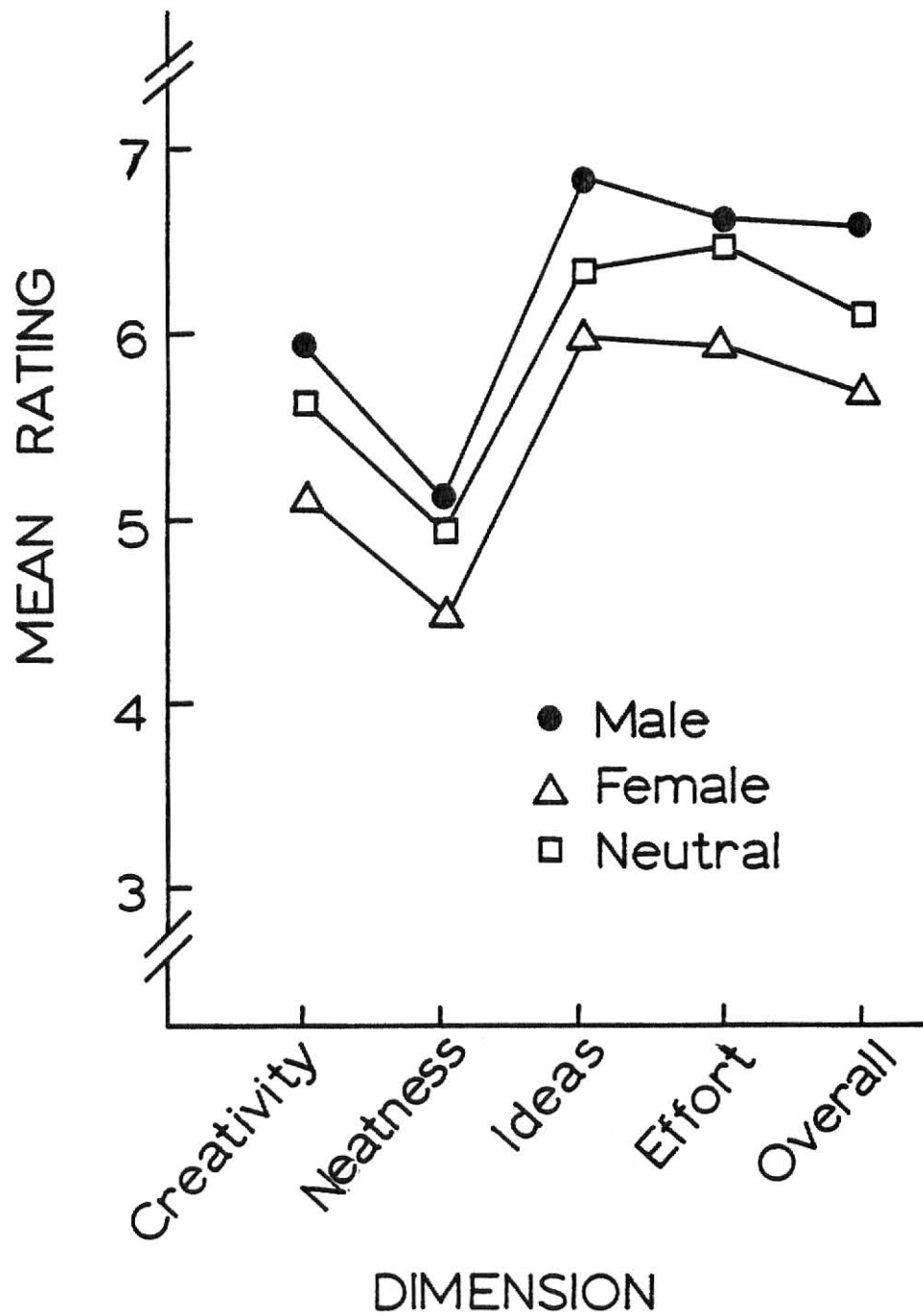
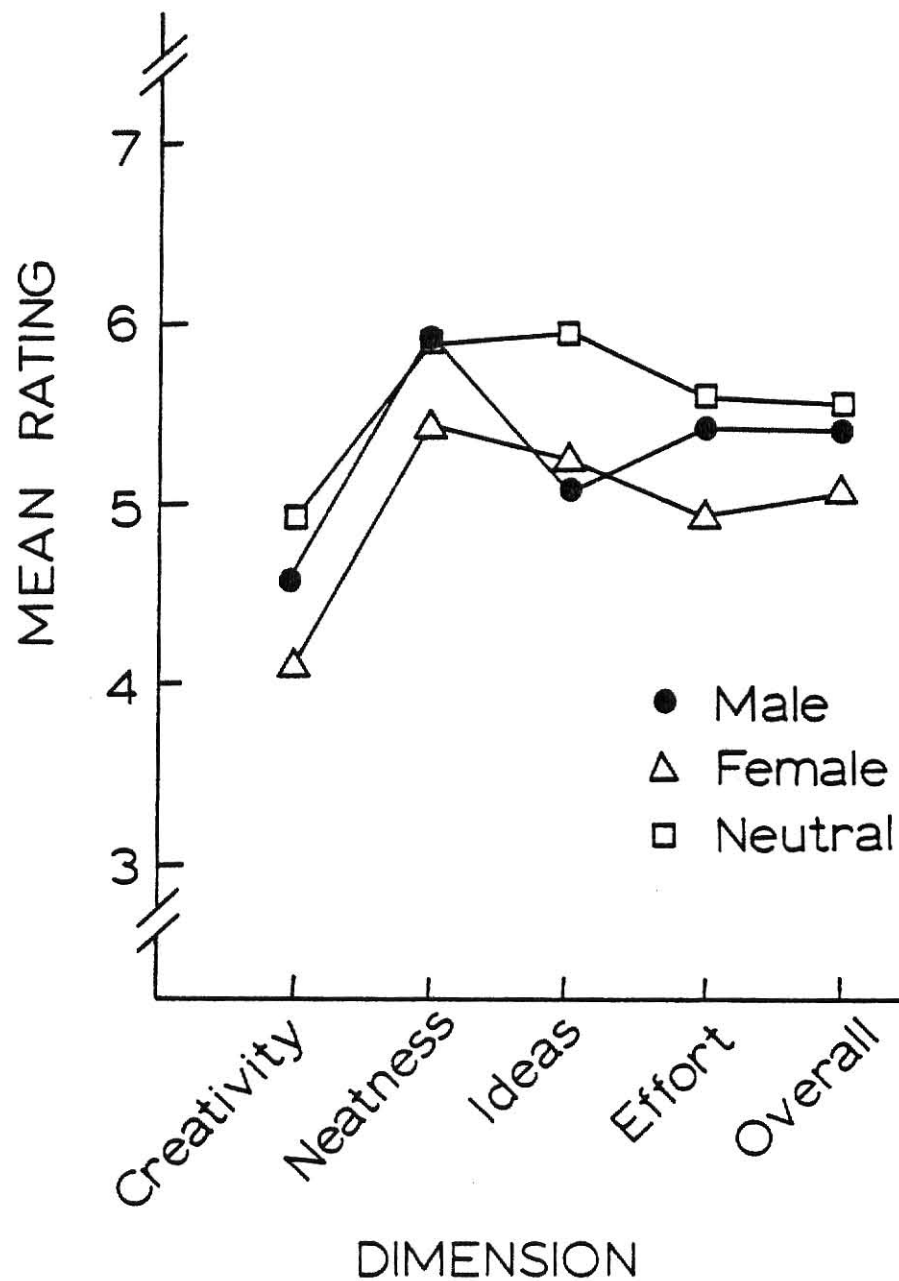


Figure 2
Mean ratings of "My Car Trip"
by sex of author

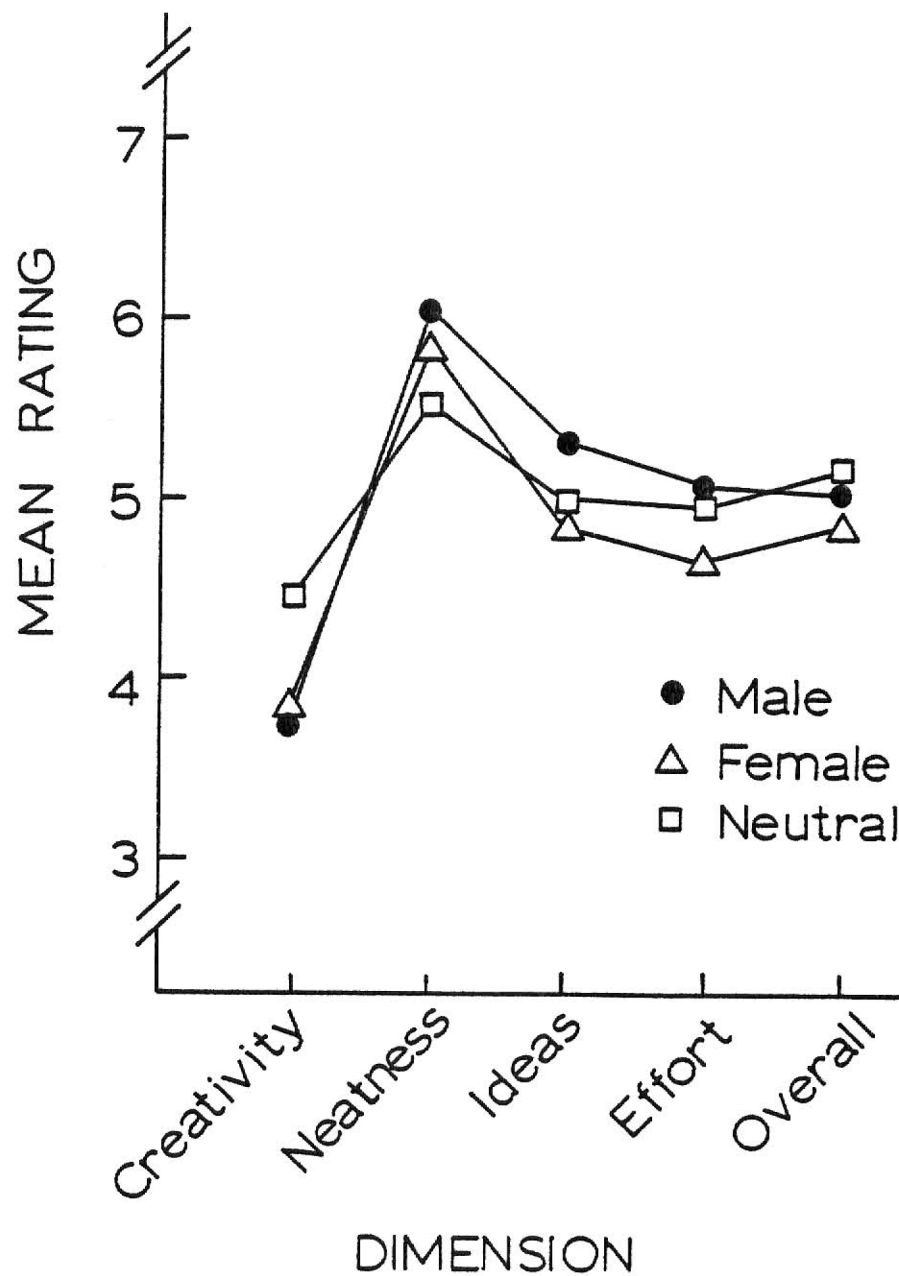


one category on which Laura was rated higher than Bill was presentation of ideas.

The mean scores for the third passage, Junior High conveying masculine sex role behaviors, are shown in Figure 3. No significant effect of author was found, $F(2,117) = .32$, $p > .10$. However, there was a significant effect of evaluation criteria, $F(4,468) = 47.06$, $p < .001$, in addition to a significant interaction between author and evaluation criteria, $F(8,468) = 2.61$, $p < .01$. The significant interaction was reflected by the control passage being rated higher than the authored passages on the creativity dimension and lower than the authored passages on neatness. Again, there was a tendency for the passage thought to be written by Pam to be rated lower than the one thought to be written by Gary; creativity was the only dimension on which Pam was rated higher than Gary.

In summary, the only passage which had a significant author effect was My Saddest Day. In that passage conveying feminine sex role behavior, Mike was actually rated higher than Julie and the control on all dimensions. The significant interaction found in Junior High was due to the differences in ratings between the control and the authored passages rather than a difference between Gary and Pam's ratings. Direct comparisons across passages must be made with caution. However, it is significant to note that on 13 out of the 15 possible scores (5 dimensions for each of the three passages), the female authored passages received lower ratings than the male authored passages (binomial test, $p < .01$).

Figure 3
Mean ratings of "Junior High"
by sex of author



Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

The purpose of the present study was to determine if teachers were sex biased when evaluating the written work of elementary school students. In addition, the study was designed to examine under what conditions bias existed. Specifically, were teachers' evaluations influenced by an interaction between the gender of the student and the sex role behavior conveyed by the student's written composition?

Conclusions

Based on the findings of this study, two major conclusions may be drawn:

1. A definite sex bias in the evaluation of students' written performance was shown to occur. Although the difference in ratings between male and female authored passages was small in certain instances, males were rated higher than females on 13 out of a possible 15 ratings.
2. Sex bias was especially apparent when the written work conveyed feminine sex role behavior (My Saddest Day). Males were rated significantly higher than females on all dimensions.

Discussion

The sex bias found in the present study was consistent with previous work which indicated that the performance of males was rated higher than the identical performance of females (Goldberg, 1968; Pheterson et al., 1971; Deaux & Taynor, 1973; Heneman, 1977). However, the higher ratings given to male students were somewhat surprising based on the data of Maccoby and Jacklin (1974) which revealed that elementary school girls are given better grades than boys. Perhaps, as discussed in the McCandless et al. (1972) study, the personality characteristics of students do affect a teacher's evaluation of their performance. In other words, in the present study teachers evaluated examples of work without being influenced by the personality or behavior of the particular student.

Female authored passages were rated higher than corresponding male authored passages on only two of the fifteen ratings, creativity (Junior High) and presentation of ideas (My Car Trip). Based on Dweck et al. (1978) one would have predicted higher ratings for girls on nonintellectual aspects of work such as effort or neatness. Yet, in the present study boys were even rated higher on these dimensions, particularly when the written work displayed feminine sex role behavior.

The equity theory could account for the higher ratings given to males on work conveying feminine sex role behavior. In the Taynor and Deaux (1973) and Bigoness (1976) studies females were rewarded more than males for identical performances

in situations that were considered masculine. Based on the nonvoluntary constraint of gender, the rewards were adjusted to achieve a sense of balance or equity. In the present study, perhaps boys were rewarded more (given higher ratings) than girls because the subject on which they were writing was considered atypical for their gender. It is almost as if to say girls are emotional and therefore it is expected that they will write about emotional experiences. Boys are not emotional and do not typically write about emotional experiences; therefore when they do, they should be rewarded by high ratings. In that way, a sense of equity is achieved.

Interestingly, the equity theory did not seem to apply to the evaluations of female work conveying male sex role behavior (Junior High). According to the theory, one would have predicted that females would be rewarded for writing about situations which were typically masculine. Perhaps the selection of atypical classes for junior high by females was interpreted as assertive, nonconforming behavior. In that case, one is reminded of the findings that females in highly assertive or nonconforming roles were least preferred by teachers and rated lower than males in intelligence (Feshbach, 1969; Etaugh & Hughes, 1975; Lao et al., 1975).

Both the higher ratings given to males on My Saddest Day and the lower ratings given to females on Junior High are inconsistent with the Bernard (1979) and the Bernard et al. (1981) studies. In their work, data indicated that high levels of

masculinity rather than femininity, regardless of the gender of the student, were rewarded. Accordingly, it would seem that males would not be rewarded for writing about an emotional reaction to death. One would also expect females to be rewarded for displaying masculine behavior in the selection of classes for junior high. Again, Bernard's studies were conducted with high school students. The attitudes and expectations regarding sex role behaviors and the bias exhibited at various grade levels seem to be quite different and must be examined further.

The issue of sex bias in evaluation of performance is extremely complex. Certainly, in this study, females seemed to be the victims of bias whether or not their work conveyed sex role behaviors that were "typical" for their gender. The equity theory was employed as a possible explanation for this bias. Yet even it was limited in its applicability. Clearly a need exists for further study in this area which would clarify other situations where sex bias might exist and the possible explanations for it.

Recommendations

The following recommendations are suggested for further research:

1. Restrict the population that is being tested to elementary school teachers (preferably upper elementary) whose teaching assignment includes the evaluation of written work. Included in the present study were secondary teachers who

rated the work of elementary school students. A more restricted population of subjects may provide slightly different results.

2. Replicate the study at different grade levels to see if the results are consistent. The passages to be evaluated should be tailored appropriately for each level.

3. Replicate the study in different parts of the country. This will help determine if the same biases or attitudes exist in other regions.

4. Conduct the study with student teachers. The present research involved teachers who had at least one year of teaching experience. By comparing the results of student teachers with experienced teachers, a more comprehensive picture may be obtained.

5. Examine carefully potentially significant variables such as female vs male teachers, school size, urban vs rural areas, number of years teaching experience, or majority vs minority responses. Again, each of these may help in specifying the conditions under which bias is most likely to exist.

The existence of sex bias in the evaluations of students' work as shown in the present study, has implications for the educational process. Prior research examining sex bias in the classroom has dealt with verbal interactions and differential treatment given to male and female students by teachers. The present study identified another area where sex bias occurs in the classroom: the evaluation of written performance. Identifying that a problem exists is an important first step.

Making other educators aware of the problem is the next vital step in effecting a change.

REFERENCES

- Adams, J. S. (1965) Inequity in social exchange. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 2). New York: Academic Press.
- Bernard, M. E. (1979) Does sex-role behavior influence the way teachers evaluate students? Journal of Educational Psychology, 71, 553-562.
- Bernard, M. E., Elsworth, F., Keefauver, L. W., & Naylor, F. D. (1981) Sex-role behavior and gender in teacher-student evaluations. Journal of Educational Psychology, 73, 681-696.
- Bigoness, W. (1976) Effect of applicant's sex, race and performance on employers' performance ratings: Some additional findings. Journal of Applied Psychology, 61, 80-84.
- Cherry, L. (1975) The preschool teacher-child dyad: Sex differences in verbal interaction. Child Development, 46, 532-535.
- Deaux, K., & Emswiller, T. (1974) Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female. Journal of Personality and Social Psychology, 29, 80-85.

- Deaux, K., & Taynor, J. (1973) Evaluation of male and female ability: Bias works two ways. Psychology Reports, 32, 261-262.
- Duval, C. M. (1980) Differential teacher grading behavior toward female students of mathematics. Journal for Research in Mathematics Education, 11, 202-214.
- Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978) Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. Developmental Psychology, 14, 268-276.
- Etaugh, C., & Hughes, V. (1975) Teacher's evaluations of sex typed behaviors in children: The role of teacher sex and school setting. Developmental Psychology, 11, 394-395.
- Feshbach, N. (1969) Student teacher preferences for elementary school pupils varying in personality characteristics. Journal of Educational Psychology, 60, 126-132.
- Goldberg, P. (1968) Are women prejudiced against women? . Trans-Action, 5, 28-30.
- Good, T. L., Sikes, J. N., & Brophy, J. E. (1973) Effects of teacher sex and student sex on classroom interactions. Journal of Educational Psychology, 65, 74-87.
- Hall, R. M., & Sandler, B. R. (1982) The classroom climate: A chilly one for women? Project on the Status and Education of WOMEN, February 1982.

- Heneman, H. G. (1977) Impact of test information and applicant sex on applicant evaluations in a selection simulation. Journal of Applied Psychology, 62, 524-526.
- Lao, R. C., Upchurch, W. H., Corwin, B. J., & Grossnickle, W. F. (1975) Biased attitudes towards females as indicated by ratings of intelligence and likeability. Psychological Reports, 37, 1315-1320.
- Leventhal, G. S., & Michaels, J. W. (1971) Locus of cause and equity motivation as determinants of reward allocation. Journal of Personality and Social Psychology, 19, 114-118.
- Maccoby, E., & Jacklin, C. (1974) The psychology of sex differences. Stanford, CA: Stanford University Press.
- McCandless, B., Roberts, A., & Starnes, T. (1972) Teachers' marks, achievement test scores, and aptitude relations with respect to social class, race, and sex. Journal of Educational Psychology, 63, 153-159.
- Pheterson, G. T., Kiesler, S. B., & Goldberg, P. A. (1971) Evaluation of the performance of women as a function of their sex, achievement and personality history. Journal of Personality and Social Psychology, 19, 114-118.
- Sadker, M., & Sadker, D. (1982) Sex equity handbook for schools. New York: Longman.
- Siegel, S. (1956) Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.

- Sternglanz, S. H., & Lyberger-Ficek, S. (1977) Sex differences in student-teacher interactions in the college classroom. Sex Roles, 3, 345-352.
- Taynor, J., & Deaux, K. (1973) When women are more deserving than men: Equity, attribution and perceived sex differences. Journal of Personality and Social Psychology, 28, 360-367.
- Taynor, J., & Deaux, K. (1975) Equity and perceived sex differences: Role behavior as defined by the task, the mode and the action. Journal of Personality and Social Psychology, 32, 381-390.

APPENDIX A

Passages Evaluated by Subjects
in Pilot Study 2 and Experiment

My Saddest Day

When I found out my grandma had died, I was so upset I ran into my room and slammed the door. I remember when I had been 7, I would sit on her lap and she would sing to me. Every time I would visit her, she would buy me a new outfit. I lay down on my bed and cried myself to sleep.

My Car Trip

From Denver to Kansas City is a long drive and the scenery is boring. No one wanted to play cards. I couldn't do my homework because the road was so bumpy. I even tried to fall asleep and couldn't. When we finally stopped for lunch I found out there was only a half hour left, so I sat and enjoyed my lunch.

Junior High

I sat down and looked over the list of classes I could take in junior high school. I definitely did not want to take chorus, but woodwork sounded good. Since I had to take science, I decided to take a science with dissection. After choosing classes, I sat down in front of the TV and watched the baseball game.

APPENDIX B

Rating Sheet for Pilot Study 2

Name of Passage _____

	shows no creativity			average creativity					extremely creative	
Creativity	1	2	3	4	5	6	7	8	9	

	messy-difficult to read				average neatness				extremely neat	
Neatness	1	2	3	4	5	6	7	8	9	

	poor				average					excellent	
Presentation of ideas	1	2	3	4	5	6	7	8	9		

	shows no effort				average effort				shows great effort	
Effort	1	2	3	4	5	6	7	8	9	

	poor				average					excellent	
Overall Evaluation	1	2	3	4	5	6	7	8	9		

	strongly masculine				neutral					strongly feminine	
Sex role behavior conveyed	1	2	3	4	5	6	7	8	9		

APPENDIX C

Cover Page from Materials Given to Subjects

The purpose of the present research is to determine how teachers evaluate the quality of written work done by 6th grade students. We would appreciate the following information:

Age: 21-30 _____ 31-40 _____ 41-50 _____ 51-60 _____ >60 _____

Sex: _____ Male _____ Female

Race: _____ Caucasian _____ Black _____ Native American
 _____ Hispanic _____ Asian American _____ Other

Total years of teaching experience: _____

Years of teaching experience at specific grade levels:

subject field if applicable

_____ K	_____ 7	_____
_____ 1	_____ 8	_____
_____ 2	_____ 9	_____
_____ 3	_____ 10	_____
_____ 4	_____ 11	_____
_____ 5	_____ 12	_____
_____ 6		

Date of last teaching assignment _____ 1982-83 _____ 1981-82
 _____ 1980-81 _____ before 1980

Grade level and subject taught _____

Type of school _____ Public _____ Private

School community _____ Rural _____ Urban _____ Suburban

Socioeconomic background of students

_____ Low _____ Medium _____ High

Was evaluating students' written work part of your past teaching assignment?

_____ Yes _____ No

APPENDIX D
Rating Sheet Used in Experiment

Name of Passage _____

Name of Student _____

	shows no creativity			average creativity				extremely creative	
Creativity	1	2	3	4	5	6	7	8	9

	messy-difficult to read			average neatness				extremely neat	
Neatness	1	2	3	4	5	6	7	8	9

	poor			average				excellent	
Presentation of ideas	1	2	3	4	5	6	7	8	9

	shows no effort			average effort				shows great effort	
Effort	1	2	3	4	5	6	7	8	9

	poor			average				excellent	
Overall Evaluation	1	2	3	4	5	6	7	8	9

SEX DISCRIMINATION IN THE EVALUATION
OF STUDENTS' WRITTEN COMPOSITIONS

by

NANCY F. KIEFER

A.B., Washington University, 1973

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Curriculum and Instruction

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1983

ABSTRACT

The objective of this study was to determine if teachers are sex biased when evaluating the written work of elementary school students. Further, are the teachers' evaluations influenced by an interaction between the gender of the student and the sex role behavior conveyed by the student's written composition?

One hundred and twenty male and female subjects were randomly selected from graduate education classes at Kansas State University. Each subject received a packet of materials containing a cover page, three passages, and three rating sheets. The cover page included personal data questions to be answered by each subject. Inside the packet the subject found three passages which were xeroxed copies of handwritten compositions done by sixth grade students. A pilot study was conducted to identify students whose handwriting was perceived as being neutral (neither feminine nor masculine) in style. The passages were further controlled for length, T-unit length, and ideation fluency. The three passages evaluated by subjects conveyed different sex role behaviors as indicated by pilot data. One passage described an emotional reaction to the death of a grandmother (female sex role behavior conveyed), another described selecting classes for junior high school (male sex role behavior

conveyed), and a third described a car trip (neutral sex role behavior conveyed). For experimental subjects ($n=80$) each of the three passages had either a male or a female name on the composition. All possible name/passage combinations were randomly distributed to the subjects. Control subjects ($n=40$) received the same passages with no student name. Three rating sheets were included in the packet of materials. For the experimental subjects, a place was designated for filling in the title of the passage and the student author. Having the subjects write down the name of the student author assured awareness of the child's gender. Control subjects' rating sheets provided only a space for the title of the passage. Subjects then rated the passages on five dimensions: creativity, neatness, effort, presentation of ideas, and overall evaluation by circling the appropriate number (1-9) on each scale.

Analysis of the passage conveying feminine sex role behavior revealed a significant author effect ($F(2,117)=5.3$, $p<.01$). Males were consistently rated higher than controls on all dimensions. Females were consistently rated lower than controls on all dimensions. On each of the other two passages (male sex role, neutral), although males were rated slightly higher on 4/5 dimensions, no significant differences were found. The data suggest that a sex bias does exist in the evaluation of students' written work. This bias is particularly evident when the work thought to be written by a male conveys feminine sex role behavior.