Undergraduate Psychology Research Methods Journal

Volume 1 | Issue 6 Article 2

2007

Gender Bias in Peer Grading Among Undergraduate Students

Elizabeth Dalton
Lindenwood University

Gender Bias in Peer Grading Among Undergraduate Students

Elizabeth Dalton

Using the observation that people refer to their own genders when talking about an unknown author, one might infer that people might relate to authors of their same sex. If this is true, this could be a gender bias as people could attribute good qualities of an unknown author to their gender. This led to an investigation of gender bias in peer grading where students were thought to attribute better grades on a paper if the author is their same sex. Participants were separated into three groups separated by the knowledge of the author's gender then asked to grade a paper. The results showed higher scores in the same sex group raising the issue of importance of anonymous peer grading.

While observing in various undergraduate classes I began to notice that when a male student was referring to a document that had no identifying characteristics, he began to refer to the document as one that was written by a man. During the same class period, and referring to the same document, a female student began to discuss the document referring to the author as a woman. Astonished by the complete ignorance that the two students had toward their references of what they perceived was the author's gender, I began to wonder whether men and women could perhaps identify better with authors that they think are men and women. This thought process led me to ponder whether sex was a defining influence on how people grade. Perhaps, men and women attribute better grades to those of the same sex as opposed to the opposite sex.

If men and women attribute better grades to those of the same sex, this could be affecting the academic careers of students. For example, if a woman was grading a paper written by a man, she might attribute a grade lower than the one she would have given to another woman. The woman might have placed a grade on a paper that would have been better or more adequate if the gender of the author were unknown. Perhaps people unknowingly engage in gender bias in order to ensure the superiority of their particular gender and to stifle this problem, steps should be taken to ensure complete anonymity of students when they are peer grading. Gender bias is considered to be "the prejudice in action or treatment against a person on the basis of their sex," ("Legal," n.d.).

There have been various studies conducted that have asked the question of gender bias in grading. In one such study, Welsh (2004) predicted that females would give better grades to authors that were men. She conducted a one-way ANOVA to find that there was no statistical significance showing that men received a better grade than women. Because there was no significance found, it gave way to the idea that perhaps Welsh should have looked at same-sex bias as opposed to opposite-sex bias. Her study led me to believe that women might have attributed better grades to other women as opposed to men.

In another study Greatorex and Ball (2004) state that, "sex and gender bias in marking should be monitored…but is unlikely to be found to an extent that affects grades (32)." Although this study did not specifically study the effects of gender bias and grading it does give way that it is present in education. They acknowledge its existence then refute its effects. In regards to a possible sex bias by professors, Manley (1933) shows that a sex bias is evident in various professors over a longitudinal study. He stated

that some professors were shown to have a bias towards college students of the same gender whereas others did not. His study was understood to identify that there is a risk of bias in grading.

Cho, Schunn, and Wilson (2006) found in their study of the validity of peer assessment that only minimal assessments were viewed by students as a valid form of grading. They found that students do not feel as though peer grading is an adequate form of academic assessment. This study suggests that students should not be prompted to participate in peer assessment as they are not in the same academic level as the teacher. Perhaps students who grade papers will be make more uneducated mistakes such as unknowingly grading with a gender bias and letting that biases affect the grade of the paper.

The current study randomly assigned participants into three groups to test the hypothesis that participants will attribute better grades to authors of the same sex. The first group was the "same sex" group where participants were told that the author of a paper is of the same sex. The second group was the "opposite sex" group where the participants were told that the author of the same paper will be of the opposite sex, and the third group will be the control group where the participants will have no knowledge of the sex of the author. Participants will grade the same paper in the same room with the same rubric in attempt to control for extraneous variables. The only variable that will differ, or will be considered the independent variable, will be the participant's knowledge of the gender of the author. The grade the participant attributes to the paper will be the dependent variable.

Method

Participants

There were a total of 74 participants tested by use of the Human Subject Pool at Lindenwood University. Participants were recruited from various social science survey classes including psychology, sociology, and anthropology at Lindenwood University. Of those 74 participants only 46 participants' data were used. The 28 other participants' data was discarded on the account that the experimenter could not test those who spoke English as their second language. Of the 46 participants, 32 were women and 14 were men. There were 19 freshmen, 12 sophomores, eight juniors, and seven seniors that participated in this study and their grade point averages ranged from 2.0 to 4.2 with the mode at 3.3. Participants were recruited from various social science survey classes including psychology, sociology, and anthropology at Lindenwood University. They were given extra credit towards those classes of which they were involved as compensation for their time and energy spent on the experiment. Although no participants dropped out of the study, there would have been no pressure if any participant decided to leave the experiment. The experimenter understood that all of the participants were to be treated equally in accordance to the ethical guidelines that were established by the American Psychological Association

Materials

The materials used consisted of a paper, a survey, and a rubric as well as the informed consent form, the list of participants, the feedback sheet, and the extra credit slips. Pens used to grade the papers were provided by the experimenter. The paper was taken from the 2007 Advanced Placement Exam with the permission of Dr. Schnellmann,

an English professor at Lindenwood University. The paper was rather mediocre in its presentation and it was slightly modified, to ensure that the grade was not easily attributed as extremely poor or well. The paper was also completely anonymous, completely leaving out any identifying characteristics to ensure that the gender, as well as the specific identity of the author could not have been known (see Appendix A for copy of paper used).

The rubric, the item the participants used to grade the paper, was provided by Dr. Nohara-LeClair, a professor of psychology at Lindenwood University. The rubric allowed participants to circle what they believed was appropriate under the areas of format, spelling and grammar, organization, clarity, and their presentation of agreements (See Appendix B for copy of rubric used). The survey was composed by the experimenter. It asked participants to state their sex, class standing at Lindenwood University, their grade point average, their first language, whether they felt as though their grading was influenced by the gender of the author, and whether they felt as though teachers' grading could be influenced by the knowledge of the author's sex (See Appendix C for copy of survey used).

Additionally, the informed consent form as well as the feedback letter was composed by the author. The list of participants and the extra credit slips were provided by the Lindenwood University Human Subject Pool. The room that was used to conduct the experiment was a part of the Psychology lab at Lindenwood University. The room used consisted of a table, three chairs, and a desk with a computer. The participants were seated at a chair on the left side of the table while The experimenter was seated at the desk with the computer.

Procedure

The experiment was constructed using a between groups design where there were three different conditions used. Each participant was randomly assigned into one of three different conditions. The first condition was referred to as the "same sex." condition. Participants were told that the sex of the author was the same as their own. The next condition was referred to as "opposite sex," where the participants were told the sex of the author was the opposite of their own. The third condition was the control where the participants were told nothing of the gender of the author.

When the participant was ushered into the room, they were first asked to read over and sign the informed consent after the experimenter explained the form. They were then asked to complete the list of participants. After this was completed, the participants were given the paper and told specific information pertaining to the condition they were randomly assigned. They were also told to use the rubric to grade the paper and they were permitted to mark on the paper as needed. After the participants had finished grading, they were then given the survey to complete. After finishing, they were then debriefed using the feedback letter as a reference. After fully discussing the feedback letter and thanking them for their participation, the participants were given the extra credit slip which was fully explained.

Results

Upon collecting the data the researcher used SPSS to analyze the research. A one-way analysis of variance was used to analyze the results. The results revealed a statistically significant effect of group assignment on grade assigned to the paper, $F_{(2, 43)} = 7.32$, $p = .002$, showing that the experimenter could reject the null hypothesis. The

results of a series of Tukey post-hoc tests revealed that the grade assigned by the same sex group (M=25.81) was significantly higher than the grade assigned by the opposite sex group (M=17.9) and the control group (M=20.93). In order to determine whether there were differences in responses between participants in the same sex group as opposed to those in the opposite sex groups in regards to their answers to the question of how much the knowledge of gender of the author of the paper effected their scoring, a Mann-Whitney U test was conducted on the ordinal data collected. The test showed no difference between the two groups.

Discussion

The experiment was designed to examine whether peers would give a better grade to those of the same sex as opposed to the opposite sex. Upon analysis, the results do support the hypothesis suggesting that there is a gender bias in peer grading. Perhaps there could have been more details placed in to action to ensure that these results were from no other confounding variables. The researcher could have had separate rubrics to be given to each participant in the conditions. For instance, if a woman was in the "same sex" group, she should have been given a rubric that stated the same information yet also stated the gender of the author, as a woman instead of only being told once by the experimenter. This would ensure that the participant would not forget the gender of the author while grading as the participants might have forgotten or not even listened to the experimenter during the process.

Another idea that might have been an influencing factor is the realization by the experimenter that the paper was truly very mediocre. This paper was not an average paper; it was actually rather poor in terms of writing quality. Various participants in all

of the conditions remarked about how the paper should not have been written for an Advanced Placement exam. This could have generated much of the lower scores although the results do show clearly that sex could have been a defining factor in the author's grading.

Despite the idea that this paper did deserve the poor grades it received, the feedback the experimenter received was uncanny. Many participants, when told of the nature and purpose of the study were very adamant in saying that they believed that they had felt as though their grades were influenced by their gender. Some participants specified that they felt as though they had benefited from taking classes with a teacher of the same gender as they felt the teacher could identify better with their work and personal problems and they received no penalties for absences or rushed work. Others actually felt as though teachers, as well as students held a bias of superiority to the opposite sex. One woman, who was in the "same sex" group actually said that if she thought the paper were written by a man, she absolutely knew that she would have been harsher when grading.

Another realization was that everyone who participated in the study did not think that their knowledge of the author's sex influenced their grading. In fact, the highest number that was circled was "3" which meant that the participant thought the knowledge of the author's gender only influenced them "somewhat" while they were grading. Yet results can be inferred that the information given to the participant was taken into account while grading. Even though the participants did not think that the information they were told about the gender of the author affected their grading, they did think that teachers' grading could be affected by the knowledge of the author's gender. Answers ranging to

"5" indicated that even though the participants did not think that they could be influenced by the author's gender, they did think that teachers' grading could. From inferring the results, one might ask the question of whether the same results could be seen when testing teachers as well.

Something that could be taken into account is the number of women who participated in the study compared to that of men. There were 32 women used as opposed to 14 men. Perhaps these results were dramatically influenced by the women's participation. Further studies should be done in effort to determine whether this bias is seen mostly in women or if this is true in both sexes. Perhaps the experimenter should have divided the groups between men and women and tested between the two genders. Obviously, it could not have been feasibly done as one could not account for the number of men and women who signed up through the Lindenwood University Human Subject Pool.

Another drawback that the experimenter experienced was the numerous participants who are described as ESL students, or students who spoke English as their second language. As mentioned earlier, 74 participants were run and 28 pieces of data were discarded. As difficult as it was to continuously test, I began to wonder if the same bias occurred in ESL students even though they are grading a paper that was not in their native language. Further studies might be done to determine whether the same bias could be evident in ESL students.

Finally, during feedback, while the researcher was telling the participants about not knowing who actually wrote the paper and they were told a specific gender for experimental purposes, many people immediately claimed that the author was most likely

a man. Even those who were told that the sex of the author was a woman admitted that now that they understood that the author was not a woman, could see the author as being male. Perhaps the researcher could have tried to find a more neutral paper but there was no possible way for me to identify the paper as written by a man or woman which was thought to ensure that there would be no experimenter bias towards the gender of the author of the paper.

Perhaps, in order to stifle this proposed problem, teachers should engage in various tasks to ensure the sex of the author's confidentiality. This might better enhance grading and produce a more accurate grade. Students should be graded on their abilities. No other characteristics should interfere with students and their academic progress. If teachers can ensure complete confidentiality of their students, they are engaging in their part to ensure an adequate grade.

References

Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of the Scaffolded peer assessment of writing from instructor and student perspectives. .*Journal of American Psychology*, *98*(4), 891-901.

Greatorex, J. & Bell, J. (2006). *Research in Education*, *71*, 25-36.

Legal Explanations. http://www.legal-explanations.com/definitions/gender-bias.htm. (November 26, 2007)

Manley, C. (1933). Sex-bias in college marking. *Journal of Higher Education*, *4*(1), 29-31.

Welsh, R. M. (2004). Gender bias: Are women biased against women. Missouri Western State University. http://clearinghouse.missouriwestern.edu/manuscripts/526.asp. (October 26, 2007)

Appendix A

Most members of society are bombarded with advertisements everyday. This will tell them many things about what they should buy or do and the smart consumer will keep in mind their own values and interests. However many will make decisions that are bad for themselves and others. Because of their powerfully persuasive nature, advertisements can present people with extra decisions and cause them to make these decisions and cause them to make these decisions without necessarily considering their own values.

Advertising is an extremely powerful method of persuasion. It has come to be so after years of research and money. That consumer research is done suggests that companies are decisive in their campaigns to sell product. Also, because the research is described as sophisticated, it is likely affective and thorough, using state-of-the-art technology and powerful thinkers. This emphasizes the decisiveness of company advertisement strategy. At the end of the first paragraph, Shaw continues to describe the size and precision of advertising campaigns. Nobody is to be left without seeing the advertisements. Advertising companies reach far and work hard to persuade consumers.

Advertising is a very powerful means of persuasion, but its main purpose is to make people aware of something. Advertising is designed to inform people. Because everyone is exposed to advertising, some use it to present good ideas or products that people really need. If people must buy something, it is best that they know who to buy it from. Advertising allows people to see what is available to purchase and decide what is best for them.

While advertising is used to inform people of important things, it is also used to make people decide on things that are bad for them and others. Sesana, (2004) provides

an account of powdered milk advertised in areas without a healthy water supply, making many children sick. Advertising has too often been used to make people buy things that will hurt them or spend money on things that they don't need.

Advertising has its good and bad aspects. It is good because it informs people of things they need to know. Unfortunately, people have used this powerful persuasive device to convince people to ignore what it best.

Appendix B

Rubric

The paper you are grading was an essay that was written for the 2007, Advanced Placement Exam. The writer was asked to write an essay developing their opinion on whether or not advertising is an adequate way to inform people or if it is used poorly. The writer was asked to defend his argument thoroughly and persuasively. Pease read the paper and grade according to the provided rubric. Please feel free to take your time and mark on the paper itself if you need to. If you have any questions let me know.

Please circle the appropriate number that you feel best reflects the paper you are grading.

Fulfillment of Assignment Format Requirements (8 points)

(e.g., length, double-spaced, typed, margins, font, font size)

0--------1---------2---------3--------4---------5--------6---------7--------8

Very Poor	Major Problems	You Can Do Better	Minor Problems	Virtually No Problems

Spelling and Grammar (8 points)

0--------1---------2---------3--------4---------5--------6---------7--------8

Very Poor	Major Problems	You Can Do Better	Minor Problems	Virtually No Problems

Organization (8 points)

0--------1---------2---------3--------4---------5--------6---------7--------8

No Organization	Scattered Thoughts	You Are Almost There	Well Organized	Very Well Organized

Clarity of Writing (8 points)

0--------1---------2---------3--------4---------5--------6---------7--------8

Unclear	Fuzzy	Cloudy	Pretty Clear	Clear As Day

Relevance and Ability to Present Convincing Arguments (8 points)

0--------1---------2---------3--------4---------5--------6---------7--------8

Not Enough Effort	Very Poor	You Can Do Better	Very Good Job	Excellent Work

Appendix C

Survey

Group letter and number _____

Please circle: Male Female

How much do you think that the information that you were told about the author's sex influenced your grading?

1 2 3 4 5

"1" being the least, "5" being the greatest

How much do you think a teacher's grading could be influenced by the knowledge of the author's sex?

1 2 3 4 5

"1" being the least, "5" being the greatest

Please circle: Freshman Sophomore Junior Senior

What is your GPA? _____

Is English your first language? _____

Author Note

Elizabeth Dalton affiliated with the Psychology Department at Lindenwood University.

All laboratory access and financial aid was afforded by the Human Subject Pool through Lindenwood University. The paper was provided by Dr. Ana Schnellmann of the English Department of Lindenwood University and the rubric and massive support was provided by Dr. Michiko Nohara-LeClair of the Psychology Department of Lindenwood University. To both of these people I hold much gratitude. Approval for this project was given by the International Review Board of Lindenwood University.

For further information concerning this study, please contact Elizabeth Dalton at edk9sm@gmail.com or Dr. Nohara-LeClair at mnohara-leclair@lindenwood.edu.