IBFR
www.theIBFR.COM

EXAMINING IF GRADING BIAS EXISTS IN A PROFESSOR OF BUSINESS COURSES: AN INDIVIDUAL STUDY

Gary F. Keller, Eastern Oregon University, LaGrande, OR

ABSTRACT

This research paper investigated if grading bias based on gender, course modality (face-to-face and online) and grade level (undergraduate vs. graduate) existed in classes the researcher taught over a three-year period. During the course of the literature review for this investigation, a meta-analysis suggested that more research was needed on a micro level regarding bias rather than reliance on results from metanalysis reports. A second goal was the extension of two studies conducted by the researcher that examined if relationships existed between the amount of time students spent on course assignments and final grades and if the quantity of time a faculty member spent grading student assignments affected final grades given. In this study grades assigned during the academic years 2018 – 2020 derived from the University's Canvas Learning Management System were analyzed. The population studied consisted of 912 students enrolled in undergraduate and graduate courses taught at a public university located in the Pacific Northwest. The main results from this examination showed statistically significant differences were found in aggregate totals between the grades of males and females and the grades of females enrolled in online and face to face courses in two of the three years of the study.

JEL: I23, J16

KEYWORDS: Discrimination, Gender Stereotypes, Natural Experiment, Biased Grading

INTRODUCTION

The subject of grading bias, raises a myriad of questions as grades have a significant bearing not only on students' self-image and future employment opportunities but also on faculty members' promotions and institutional reputations. Ascertaining the causes of grading biases at a university level has been the subject of numerous, meta-analysis studies (primarily European institutions) often attempting to support/challenge the validity of anonymous grading schemes. However, a gap in the scholarly literature was discovered. Scant attention has been paid to studying individual faculty members grading tendencies rather than on an institutional level leading one researcher (Keyser, 2018) to recommend grading bias research should be conducted by individual faculty members. The chief goal of this exploratory research project was to contribute new research to the subject domain by examining courses taught by this faculty member during the academic years 2018 – 2020 to determine, to what extent if any, grading bias existed. Secondary research goals were: a) to learn if there were statistically significant differences between the final course grades for students enrolled in online and face to face courses; b) to inquire if differences exist between the final course grades of males and females and; c) to find out if disparities exist between the final course grades of undergraduate and graduate students. A review of the scholarly literature regarding grading bias is presented, an explanation of the methodology used in this research project is described, the results of statistical tests conducted to determine if grading bias existed in courses taught by the researcher are shown and recommendations for future grading bias research on the part of individual faculty members, institutions and accrediting agencies are made.

Numerous definitions exist for bias ranging from conscious to unconscious partiality. Research bias is defined as "any tendency which prevents unprejudiced consideration of a question. In research, bias occurs when "systematic error [is] introduced into sampling or testing by selecting or encouraging one outcome or answer over others" (Pannucci & Wilkins, 2011, para. 3). In grading students work, bias may be considered "as a technical term most often refers to a characteristic of tests that present advantage or disadvantage to a particular subgroup (e.g., by gender or ethnicity) (Nitko, 2004; Popham, 2005 as cited in Hardré, 2014, p. 1). In common usage, bias is "a particular tendency, trend, inclination, feeling, or opinion, especially one that is preconceived or unreasoned" (Dictionary.com, 2021). It is assumed that faculty members act with personal and professional integrity. Institutions of higher learning generally do not require faculty members to rigorously explore their grading trends other than grade distribution reports contained in promotion portfolios (which merely exhibit how many As, Bs, Cs, etc. over a period of time are assigned). Adjunct faculty members, who account for an increasing percentage of university faculty, are rarely, if ever, included in grade bias studies. While many grading tools, such as rubrics embedded in learning management systems, are available and helpful, they merely assist how one can justify assignment of grades but do not enable a faculty member to locate if grading biases exist. There appears to be few if any organizational efforts, requirements or positive individual incentives for faculty members to rigorously examine if grading biases exists in their courses. Student grade appeals and complaints, which occur on an ad hoc basis, are the generally accepted process to at least raise the question of grading unfairness. However, the grade appeal is a singular, after the fact process to raise the question of a faculty member's grading predispositions rather than a proactive, ongoing, quantitative improvement method.

Exploring grading bias and its impact on faculty and student performance is the focus of this study. It evolved from an exploration of the topic which indicated the existence of large-scale bias research projects and meta-analysis but a scant number of investigations conducted by individual faculty members. A second motivation was from two previous investigations the author conducted that sought to determine if a relationship existed between the amount of time students and this faculty member spent on the institution's LMS and final course grades. Part of the data analysis included in this study comparing face-to-face students with online students, undergraduate and graduates and male/female contrasts. One variable of this author's previous two studies that was not explored was if biases may have influenced the way this professor allocated student grades. The same subject groupings from the previous two studies were used in this research project to ascertain if grading bias may exist in courses taught by this faculty member over a three-year time period. A review of the scholarly literature encompassing the domains of existing studies regarding grading bias on a university level, the role that gender may play in grading bias and techniques that sought to decrease/eliminate grading bias are discussed followed by a description of the methodology used in this study, the results of the data analysis and recommendations for future research are offered.

LITERATURE REVIEW

Three scholarly spheres were examined in this exploratory research topic. The first domain involved locating studies that examined grading bias on a university level. A consistent theme in the literature was large-scale studies and many were European based. Malouff and Thorsteinsson (2016) conducted a meta-analysis of 23 analyses of 20 experimental research studies on populations ranging from grade school children to university students. The authors remarked that "bias can occur in subjective grading when graders are aware of irrelevant information about the students" (para.1). One important limitation noted by Malouff and Thorsteinsson was meta-analysis studies "did not examine grading of student work by the actual teachers of the students" (2016, para. 25). Malouff, Emmerton and Schutte (2013) explored the *halo effect* on grading. A cross disciplinary group of 159 professors and teaching assistants was tasked to first grade oral presentations and then written presentations of the same students. Results of this research indicated that the faculty/TAs gave significantly higher scores to written work following the better oral

presentation than following a poor demonstration displaying a halo effect. Malouff, et al. recommended that keeping students anonymous helps prevent bias in grading.

Another massive study by Hinton and Higson (2017) involved a study of more than 30,000 university students over a 12-year period to primarily ascertain if grading assignments namelessly would reduce performance differences between a variety of student classifications (ex. gender, ethnicity, social status, etc.). While Hinton and Higson concluded that anonymous marking had a negligible effect in reducing performance disparities, they urged future researchers to understand performance differences in terms of the *mechanisms* by which performance differences manifest at the group level…and it is far from easy in practical terms to obtain these kinds of data in large quantities" (para. 60). Bygen (2019) reviewed data from five undergraduate courses at Stockholm University, from 2005 and 2013 to see if examination results were biased based on the *foreignness* of one's name and gender. Results showed that grades were not biased based on these two factors. The above studies tended to be centered on utilizing large data sets to assess if grading bias existed. A study by Keyser (2018) took a different approach. The researcher centered his research on determining if course grades differed between on-campus and distance (online) students taking the same courses over a one-year period; other elements such as gender, age, etc. were also examined. Two significant differences, class standing and GPA, accounted for disparities in grades. Keyser's study was an interesting approach because it examined a single professor seeking to determine if grades may have been influenced by the medium of instruction. Additionally, Keyser recommended that professors throughout the university where he teaches should conduct studies similar to his for comparative evaluation and later expand that research to additional universities.

The second domain reviewed was the role that gender may play in grading bias. Exploring the part gender may play in university admissions, academic and social life is an important issue as American institutions continue the long effort to improve equity and inclusion in university life. Two key factors to enable fairness and inclusiveness to succeed is creating an environment that provides a transparent and reliable admission process but also supports students' educational pathways once they are admitted. Sometimes the best intentions obscure rather than clarify high minded admission goals. One study (Breda & Ly, 2015) revealed that in entrance exams of a French higher education institution, the appraisal favored females in traditionally male-dominated subjects (e.g., math, philosophy) and the reverse in customarily female-dominated areas (e.g., literature, biology). A micro study (n = 12 students) was conducted to investigate if student gender influenced feedback. The context for this study was to test the validity of those who purport that anonymity is the optimal solution to grading bias. The results of the study indicated that gender had little to no effect and thus contested the merits of the anonymity advocates.

Jansson and Tyrefors (2018) affirmed that there are few studies investigating grading bias at the university level. In their study (conducted in Sweden) evidence was found that "TAs correcting exams at the university favor students of their own gender. However, the size of the in-group bias was only approximately 20 % of the total effect. Interestingly, both the in-group bias and the general bias disappear when exams are graded anonymously" (p. 21). Krawczyk (2018) reported the results of a study on grades awarded for bachelor and master theses at a large Polish university. The research's purpose was to detect gender or physical attractiveness bias. Approximately 15,000 students were included in the study. The conclusions noted that some evidence existed to indicate that females received relatively high grades from advisors and no evidence of influence of physical attractiveness. Additionally, Krawczk noted "gender seems to play some role, with male students getting relatively higher grades from referees and females from advisors. This is consistent with the hypothesis of males being perceived as more competent but less likeable, and may thus be a manifestation of a bias… It does suggest that grading of term papers, which are only read by a single evaluator who knows the student personally, may be biased against male students" (p. 158). The third scholarly literature field examined focused on techniques that sought to decrease/eliminate grading bias. The French, Swedish and Polish studies in particular raise a healthy question; do universities regularly examine grade distributions based on gender, composition of the

student body, modality (i.e., face-to-face or online), and course level (freshman – senior, undergraduate – graduate)? Other queries derived from the literature review include how many faculty members regularly examine grade distributions to provide insights into their grading patterns and practices and should regional and professional accreditation agencies require grade distribution analysis as part of their endorsement process? Finally, the literature review showed that queries into the grading bias question appeared to be singular events demonstrating the need for university leadership to institute regular, consistent and quantitative research into the issue. The literature review contributed to this research project to show the problem with analyzing aggregate data rather than carefully segmentation. The results from this study indicated that some grading bias based on gender, modality or grade level existed in courses taught by this faculty member from 2018-2020.

METHODOLOGY

The chief goal of this exploratory research project was to determine if grading bias existed in courses taught by this faculty member during academic years 2018 - 2020. Secondary research goals were: a) to determine if there were statistically significant differences between the final course grades for students enrolled in online and face to face courses; b) to what extent did differences exist between the final course grades of males and females; c) to what extent did differences exist between the final course grades of undergraduate and graduate students.

This investigation encompassed 51 College of Business courses (12 face-to-face - 24%, 39 - 76% online) taught by this researcher at a small, liberal arts, rural public university located in the Pacific Northwest from the start of the Winter 2018 term to the completion of the Fall 2020 term). The total population consisted of 912 students (430 males - 47%, 482 females - 53%). Of the 912 students 780 (86%) were undergraduates, 132 (13%) graduates; 236 (26%) participated in face-to-face courses; 676 (74%) were enrolled in online courses. 116 (88%) graduate students were enrolled in online courses and 16 (12%) choose face to face courses. It should be noted that enrollment in online courses dramatically enlarged beginning in the Spring term of 2019 due to the Covid-19 pandemic. The university essentially transferred all courses to an online format; however, in the Fall 2020 term, the institution reopened with a variety of restrictions. The pandemic's impact on this faculty member's teaching load did not have a material effect on the size of the research population.

Course assignments were generally composed of quizzes, discussions, essays and a final "capstone" assignment. Data for this research was derived from the University's Canvas Learning Management System (LMS). Individual and final course grades were recorded in the Canvas LMS and recovered by the researcher. The University utilizes a quarter course scheduling system; each term is composed of 10 weeks of instruction and one week allocated for final exams. A single factor ANOVA test was used to determine differences between subjects in this study. A student t-test was used to explore differences resulting from the ANOVA analysis. A significance value of 0.05 was used to determine whether to accept or reject the null hypothesis in each hypothesis analyzed (the use of ** was used to designate significance at the 0.05 level in tables). To analyze the data, a grade point average number was assigned to each letter grade in the data set. For example, 4.0 represented "A", 3.67 represented "A-", 3.33 represented "B+", 3.0 represented "B" and so on. Finally, the data was organized by gender for each course in order to examine the hypotheses. The main research questions for this exploratory research are listed in Table 1.

Table 1: Research Questions

RQ 1	To what extent if any are there statistically significant differences between the grades of male students and female students enrolled in courses taught by this faculty member between 2018 - 2020?
RQ2	To what extent if any are there statistically significant differences between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online courses taught by this faculty member between 2018 – 2020?
RQ3	To what extent if any are there statistically significant differences between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020?
RQ4	To what extent if any are there statistically significant differences between the grades of male undergraduate face-to-face students and male online undergraduate college students in courses taught by this faculty member between 2018 - 2020?
RQ5	To what extent if any are there statistically significant differences between the grades of female undergraduate face-to-face students and female online undergraduate students in courses taught by this faculty member between 2018 – 2020?
RQ6	To what extent if any are there statistically significant differences between the grades of face-to-face undergraduate students and online undergraduate students in courses taught by this faculty member between 2018 - 2020?

Table 1 shows the main research questions that guided this study.

RESULTS AND DISCUSSION

A single factor ANOVA test was used to determine differences between the subjects in this study. A student t-test was used to explore differences if the ANOVA results were significant. A significance value of 0.05 was used to determine whether to accept or reject the null hypothesis in each hypothesis set.

RQ 1: To what extent if any are there statistically significant differences between the grades of male students and female students enrolled in courses taught by this faculty member between 2018 - 2020? (**N** = 912, 430 males, 482 females)

Hypothesis: H1o: There are no statistically significant differences between the grades of male and female students enrolled in courses taught by this faculty member between 2018 – 2020.

Result: As depicted in Table 2, statistically significant differences were found between the grades of male and female students enrolled in courses taught by this faculty member between 2018 – 2020.

Table 2: Grade Comparison – Male/Females 2018 – 2020

Hypothesis-Year	T-Critical	T-Statistic	P-Value	Decision
(Male vs. Female 2018 - 2020)	1.962	2.0467	0.0041	Reject Null

Statistically significant differences were found between the grades of male and female students enrolled in courses taught by this faculty member between 2018 – 2020. The chief reasons for the variances may be due to combining all types of courses (undergraduate/graduate), modalities (face to face and online, disparities between the number of males (430) and females (482) and effect of aggregating dissimilar groups that indicated grading bias on a macro level while the micro-outcomes indicate little if any grading bias.

RQ 2: To what extent if any are there statistically significant differences between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online courses taught by this faculty member between 2018 – 2020? (N= 69, 11 Face to Face, 58 Online).

Hypothesis H2o: There are no statistically significant differences between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online courses taught by this faculty member between 2018 – 2020.

Result: Data displayed in Table 3 shows that no statistically significant differences were found between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online courses taught by this faculty member between 2018-2020.

RQ 3: To what extent if any are there statistically significant differences between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020? (*N=6 Face to Face, 57 Online)*.

Hypothesis H3o: There is no statistically significant differences between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020.

Result: Data displayed in Table 3 shows that no statistically significant differences were detected between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020. Female graduate students enrolled in face-to-face courses earned higher grades than female graduate students enrolled in online courses. A key reason for the disparity was very low enrolled face to face courses compared to online.

Table 3: Grade Comparison – Male/Female Graduate Students 2018 – 2020

	T-Critical	T-Statistic	P-Value	Decision
Panel A: Male Graduate Face to Face/Male Graduate Online 2018 – 2020				
	2.07	0.090	0.9288	Accept Null
Panel B: Female Graduate Face to Face/Female Graduate Online 2019 – 2020				
	1.67	1.76**	0.042	Reject Null

*Note: ** was used to designate significance at the 0.05 level. The results of the analysis in Table 3 shows that no statistically significant differences were found between the grades of male graduate students enrolled in face-to-face courses and male graduate students enrolled in online and no statistically significant differences were detected between the grades of female graduate students enrolled in face-to-face courses and female graduate students enrolled in online courses taught by this faculty member between 2018 – 2020. Female graduate students enrolled in face-to-face courses earned higher grades than female graduate students enrolled in online courses due to very low enrolled face to face courses compared to online.*

RQ 4: To what extent if any are there statistically significant differences between the grades of male undergraduate face-to-face students and male online undergraduate college students in courses taught by this faculty member between 2018 - 2020?

Hypothesis H4o: There are no statistically significant differences between the grades of male undergraduate face-to-face students and the grades of undergraduate male online students in courses taught by this faculty member between 2018 - 2020. (Face to Face N = 118 – Online N = 256)

Result: Table 4 below shows that no statistically significant differences were found between the grades of male undergraduate face-to-face students and the grades of undergraduate male online students in courses taught by this faculty member between 2018 - 2020.

RQ 5: To what extent if any are there statistically significant differences between the grades of female undergraduate face-to-face students and female online undergraduate students in courses taught by this faculty member between 2018 – 2020?

Hypothesis H5o: There are no statistically significant differences between the grades of female undergraduate face to face students and undergraduate female online students in courses taught by this faculty member between 2018 – 2020. (*Face to Face N = 74 – Online N = 346)*

Result: Table 4 below shows that no statistically significant differences were found between the grades of male and female undergraduate regardless of modality in courses taught by this faculty member from 2018 - 2020. However statistically significant differences were found between the grades of females

enrolled in undergraduate online and face to face courses taught by this faculty member between 2019 - 2020. The disparity could be attributed to the types of courses (modality) assigned to the faculty member and major decreases in enrollment in face-to-face courses due to the Covid-19 pandemic.

Table 4: Grade Comparisons – Male/Female Undergraduate Students 2018 – 2020

	T-Critical	T-Statistic	P-Value	Decision
Panel A: Undergraduate Male Face-to-Face Students and Undergraduate Online Students				
(Male OL vs. FF 2018)	1.9812	-0.1993	0.8424	Accept Null
(Male OL vs. FF 2019)	1.9803	-1.7756	0.0784	Accept Null
(Male OL vs. FF 2020)	1.9774	0.2540	0.7999	Accept Null
Panel B: Female Face- to-Face Students and Undergraduate Online Students				
(Female OL vs. FF 2018)	1.9757	-1.3627	0.1749	Accept Null
(Female OL vs. FF 2019)	1.9808	-3.0185	0.0031	Reject Null
(Female OL vs. FF 2020)	1.9762	-2.3235	0.0215	Reject Null

Table 4 above shows that no statistically significant differences were found between the grades of male and female undergraduate regardless of modality in courses taught by this faculty member between 2018 - 2020. However statistically significant differences were found between the grades of females enrolled in undergraduate online and face to face courses taught by this faculty member between 2019 - 2020. The disparity could be attributed to the types of courses (modality) assigned to the faculty member and major decreases in enrollment in face-to-face courses due to the Covid-19 pandemic.

RQ 6: To what extent if any are there statistically significant differences between the grades of face-to-face undergraduate students and online undergraduate students in courses taught by this faculty member between 2018 - 2020?

Hypothesis: H6o: There are no statistically significant differences between the grades of face-to-face undergraduate students and online undergraduate students in courses taught by this faculty member between 2018 – 2020.

Result: Table 5 below illustrates that in general no statistically significant differences were found between the grades of undergraduate face-to-face students and the grades of undergraduate face to face online students in courses taught by this faculty member in 2018 and 2020. The primary reason for the disparity between grades between U/G and G online/face to face was attributed to the near termination of face-to-face courses due to the Covid-19 pandemic resulting in a reject the null decision in 2019. It should be noted that while the decision was to accept the null in year 2018, the p-value of 0.12 was trending towards significance as was the p-value of 0.14 in 2020.

Table 5: Grade Comparison Between Face-to-Face Undergraduate Students and Online Undergraduate Students 2018-2020 (N = 806, 285 Face-to-Face, 521 Online)

Hypotheses - Year	T-Critical	T-Statistic	P-Value	Decision
U/G Online v. U/G Face to Face 2018	1.97	1.56	0.12	Accept Null
G Online v. U/G Face to Face 2019	1.97	3.02	0.0009**	Reject Null
U/G Online v. U/G Face to Face 2020	1.97	0.91	0.14	Accept Null

*Note: ** was used to designate significance at the 0.05 level. The data analysis in Table 5 above illustrates that no statistically significant differences were found in the aggregate between the grades of undergraduate face-to-face students and the grades of undergraduate face to face online students in courses taught by this faculty member in 2018 and 2020. The primary reason for the disparity between grades between U/G and G online/face to face was attributed to the near termination of face-to-face courses due to the Covid-19 pandemic resulting in a reject the null decision in 2019. It should be noted that while the decision was to accept the null in year 2018, the p-value of 0.12 was trending towards significance as was the p-value of 0.14 in 2020.*

CONCLUSION

The chief goal of this exploratory research project was to determine if grading bias existed in courses taught by this faculty member during academic years 2018 - 2020. Six research questions were posed to respond to the research questions. A single factor ANOVA test was used to determine differences between subjects in this study. A student t-test was used to explore differences. A significance value of 0.05 was used to determine whether to accept or reject the null hypothesis in each hypothesis. To analyze the data, a grade point average number was assigned to each letter grade in the data set. Finally, the data was organized by gender for each course to examine the hypotheses. Of the six research questions, segmented findings resulted in accepting 7 null hypotheses that no grading biases between subjects existed. There were 5 instances where the null hypothesis was rejected indicating a disparity between the aggregate grades of male and female students and 4 occurrences when the subject populations were analyzed in detail in courses taught by this faculty member between 2019 and 2020. Two contributing reasons for bias could be attributed; a) the types of courses (modalities) assigned to the faculty member and changes in enrollments by modalities caused by the Covid-19 pandemic.

The implications of this work for educators are; a) further basic study needs to be done by individual faculty members as a matter of post course analysis about grading tendencies recorded in the institution's LMS or other grade records; b) a standard grade reporting methodology needs to be developed to aggregate individual faculty members grading outcomes, ex. by course segmented by genders, face-to-face, grade level, online, etc. for use by individual departments, programs, colleges, schools, etc. to enable higher level evaluation to evaluate grading trends at individual institutions; c) accreditation agencies should consider requiring reporting of grading changes and movements and actions taken by institutions to ensure that students are evaluated by faculty members utilizing fair, effective, translucent and consistent grading practices and; d) the research population (faculty and students) should be enlarged, to capture a multiplicity of populations, courses and disciplines to create a reference point from which further research can be conducted to foster improvements regarding faculty grading practices.

There were three limitations to this study. The first was gender which was presumedly based on interpersonal interactions as participants did not self-report their gender. Gender was assumed based on personal interactions and presentations. Gender for online students was grounded on commonly accepted name conventions. However; if gender identity was not clear from interactions or easily recognized from the individual's name, the participant's data was excluded from the analysis. The second limitation was the size of the subject population. The study involved only courses taught by this researcher during three calendar years and as a result did not reflect a representative sample of the entire student population of the University or College of Business. A third restraint was class size, mix of types of assignments (example an automatically graded quiz, discussions, etc.) and level (undergraduate compared to graduate level) influenced the faculty member's grading expectations. Some courses such as the BA 321 online had more quizzes and fewer discussions than their face-to-face counterparts which did not have any quizzes and more discussions. Four suggestions for future research on this topic are suggested; a) further basic study needs to be done by individual faculty members as a matter of post course analysis about grading tendencies recorded in the institution's learning management system or other grade records; b) a standard grade reporting methodology needs to be developed to aggregate individual faculty members grading outcomes, ex. by course segmented by genders, face-to-face, grade level, online, etc. for use by individual departments, programs, colleges, schools, etc. to enable higher level evaluation to evaluate grading trends at individual institutions; c) accreditation agencies should consider requiring reporting of significant grading changes and actions taken by institutions to ensure that students are evaluated by faculty members utilizing fair, effective, translucent and consistent grading practices and; d) the research population (faculty and students) should be enlarged to capture a wider range of course subjects and modalities.

REFERENCES

Breda, T., & Ly, S. T. (2015) "Professors in Core Science Fields are not Always Biased Against Women: Evidence from France" *American Economic Journal: Applied Economics*, *vol. 7*(4), p. 53–75. Retrieved June 3,2021 from: https://doi.org/10.1257/app.20140022

Bygren, M. (2019) "Biased Grades? Changes in Grading after a Blinding of Examinations Reform," *Assessment and Evaluation in Higher Education*, vol. *45*(2), p. 292–303. Retrieved May 30,2021 from: https://doi.org/10.1080/02602938.2019.1638885

Dictionary.com (2021) Retrieved June 19,2021 from: https://www.dictionary.com/browse/bias

Hardré, P. L. (2014) "Checked Your Bias Lately? Reasons and Strategies for Rural Teachers to Self-Assess for Grading Bias," *Rural Educator*, *vol. 35*(2), Retrieved May 15, 2021 from https://doi.org/10.35608/ruraled.v35i2.352

Hinton, D. P., & Higson, H. (2017) "A Large-Scale Examination of the Effectiveness of Anonymous Marking in Reducing Group Performance Differences in Higher Education Assessment," *PLOS ONE*, *vol. 12*(8), Retrieved July 7, 2021 from https://doi.org/10.1371/journal.pone.0182711

Jansson, J., & Tyrefors, B. (2018) "Gender Grading Bias at Stockholm University: Quasi-Experimental Evidence from an Anonymous Grading Reform," (SSRN Scholarly Paper ID 3334460). Retrieved June 14, 2021 from Social Science Research Network, https://papers.ssrn.com/abstract=3334460

Keyser, R. S. (2018) "Analyzing Course Grades in a Converged Classroom Environment--A Follow-up Study," *Journal of Higher Education Theory & Practice*, *vol. 18*(4), p. 102–113. Retrieved August 6, 2021 from http://search.ebscohost.com.access.library.eou.edu/login.aspx?direct=true&db=eft&AN=132449553&site=ehost-live

Krawczyk, M. (2018). Do gender and physical attractiveness affect college grades? Assessment & Evaluation in Higher Education, *43*(1), p. 151–161. Retrieved August 31, 2021 from http://search.ebscohost.com.access.library.eou.edu/login.aspx?direct=true&db=eft&AN=126358822&site=ehost-live

Malouff, J. M., Emmerton, A. J., & Schutte, N. S. (2013) "The Risk of a Halo Bias as a Reason to Keep Students Anonymous During Grading, *Teaching of Psychology*, *vol. 40*(3), p. 233–237.

Malouff, J.M., Thorsteinsson, E.B. (2016) "Bias in Grading: A Meta-Analysis of Experimental Research Findings. Retrieved September 2, 2021 from https://doi-org.access.library.eou.edu/10.1177/0004944116664618

Nitko, A.J. (2004) Educational Assessment of Students. 4th Edition, Pearson Education, Inc., Upper Saddle River, NJ. Retrieved on June 22, 2021 from https://www.scirp.org/(S(351jmbntvnsjt1aadkposzje))/reference/referencespapers.aspx?referenceid=2411143

Pannucci, C. & Wilkins, E. (2011) "Identifying and Avoiding Bias in Research," Retrieved on August 18, 2021 from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2917255/#:~:text=I.,or%20answer%20over%20others%E2%80%9D%207.

ACKNOWLEDGEMENT

BIOGRAPHY

Dr. Keller teaches management courses in the College of Business at Eastern Oregon University. He also teaches for Monarch Business School – Switzerland.