

FACULTY PERCEPTIONS OF PROJECT ESSAY GRADE (PEG) SOFTWARE

Michael J. Curran, Strayer University, michael.curran@strayer.edu
Peter Draus, Robert Morris University, draus@rmu.edu
George Maruschock, Strayer University, george.maruschock@stryaer.edu
Tim Maier, Robert Morris University, tmaier@rmu.edu

ABSTRACT

The development of increasingly sophisticated software designed to both grade and provide feedback to student written work, called Project Essay Grade (PEG) software, has created significant issues within higher education. Supporters of this technology cite virtually instant grading and feedback as a tremendous upgrade for students. They also cite several studies that indicate no significant differences in actual grades generated by the software and those developed by actual faculty. However, opponents contend that this technology may have adverse effects on the student-teacher relationship that are yet to be fully understood and they also maintain that moving towards a computer based grading system for written work invites the possibility of students 'hacking' the software and designing essays not for learning but merely to 'beat the algorithm'. Studies indicate that undergraduate students may be generally more accepting of this technology as long as the grades meet with expectations. The willingness of faculty to accept this technology remains a critical factor in the process of using PEG software on a large scale. This study surveyed 165 faculty at 3 different universities and revealed that faculty perceptions generally fell into three overarching categories: time constraints, professional/university expectations, and issues of fairness and cheating. Issues of fairness and cheating emerged as the most significant perceptual factor for faculty. ANOVA analysis further revealed that the most significant variable in willingness to accept PEG software was whether or not the faculty taught primarily online or on campus. Overall, faculty had a strong preference not to use PEG software to grade student written work.

Keywords: Project Essay Grade (PEG) software, technology acceptance, college faculty, faculty perceptions, ANOVA, Principal Component Analysis

INTRODUCTION

There has been significant growth and technological developments in the area of project essay grade (PEG) software. Research indicates that increasing levels of sophistication and 'learning' on the part of this software may make it a valuable tool in grading student writing [18]. Among the many issues that this technology addresses include increased speed in grading and feedback, a lack of bias on the part of the grader, and the ability to process large amounts of student written work [19,20,28]. Additionally, there is evidence that students would be accepting of moving the grading process of written work from faculty to software [5]. Students in one study indicated that issues of bias and slow return on graded material were significant in many cases to justify the use of computer software, especially when it can be demonstrated that the grade was commensurate with student expectations for written work [5].

While it seems that students are becoming increasingly more comfortable with the idea of a computer program grading their written work, the issue of faculty acceptance remains unclear. Many faculty view grading as a central function of teaching and the potential role of PEG software within that framework was the primary purpose of this study [9,16, 21]. The increased use of PEG software to grade students work has become an increasingly important issue in higher education. There are conflicting views currently on this issue. One side of the argument contends that both students and faculty benefit from this technology. Students benefit by having access to rapid grades and feedback, which are consistently cited as sources of concern by students. Students also may be able to use PEG software as a tool in revising and editing written work and therefore be able to submit to faculty more refined and well developed essays [10,21,22,23]. Supporters of PEG software also contend that the increasing sophistication of the programs more closely mimics faculty grading and that overall grades and feedback are statistically comparable to having actual faculty grades [10,17,29]. Supporters also note that faculty derive benefit from this software as well. Studies indicate that grading of essays is consistently one of the issues that concern faculty, both in terms of time

involvement and as a tool of assessment [1,32,34]. PEG software would allow faculty greater time to devote to other aspects of teaching, including research, publication, and spending time developing lessons.

Opponents of PEG software contend that this technology will have a negative impact on the faculty and cause damage to the impressions that both students and administrators have towards faculty [17]. If faculty expertise in evaluation can be replaced by technology, some argue, it reduces the faculty role to one of mere information dissemination, which can also be done with technology [25,31]. It also reduces the mentoring relationship that faculty have with students since grading and assessment are primary ways in which faculty develop students. Additionally, some contend that removing the 'human element' from the grading process invites a higher likelihood for cheating and that technology is eventually 'figured out' by students and that can lead to formulaic essays designed not for learning and enrichment, but merely designed to 'beat the grading algorithms' of the machine [3,11].

The purpose of this study was to discover perceptions that faculty had regarding the use of PEG software and to assess the influence that these perceptions had on the degree to which faculty were willing to accept this technology in their classes. The growing use of this technology and the possible future ramifications of PEG software make this a viable area of study. The use of this technology in both MOOCs (Massive Open Online Courses) and traditional and online classes has the potential to be a significant change in the student-teacher relationship. Among the most critical factors in this process will be the degree to which faculty both understand and accept this technology within their current faculty role. Therefore, this study evaluated faculty perceptions on the use of PEG software as an important step in understanding the changing nature of this relationship.

LITERATURE REVIEW

Faculty and Grades

Faculty grading of student work has been a well-documented issue [1,3,4,12,33]. There has also been significant research on factors that influence faculty perceptions on the value, efficacy, and importance of grading student written work [6,11]. Several major issues emerge for faculty when grading student written work. These include time commitment needed to grade, value of written work to student learning, and level of feedback necessary to justify grades [1,12,27]. Other considerations for faculty include fairness of grading, grade inflation concerns and the degree to which grading impacts a faculty members ability to gain tenure and promotions [13,14,15]. Several studies also indicate that faculty are increasingly concerned with the balancing grading of written work within the overall framework of student assessment. There is much conflicting information on how faculty perceive the grading of student written work and the value that it has in overall assessment on a student's intellectual level [2,7,8,26]. However, faculty perceptions on the nature of grading were a central aspect of this research study

The Emergence of Project Essay Grade Software

The development of PEG software has emerged as a potential solution to many of the issues expressed by students in regard to grading of written work [18,20]. These concerns include bias, timeliness, and the nature of the feedback. This software which uses sophisticated algorithms and a learning platform claims to be able to simulate faculty responses for both grading and feedback and can do so in a fraction of the time that it usually takes faculty [30]. Several universities and the Educational Testing Service (ETS) have begun to look at the impact this technology can have [17]. Several studies indicate that this technology can benefit students, save time for faculty, and be a cost saving mechanism for universities, especially those with large online components [20,28]. Additionally, one study indicates that students, especially undergraduates, are accepting of this technology [5]. However, as universities consider the merits of adopting this technology, gaining the acceptance of faculty becomes a crucial element in the process. The purpose of this study was to understand how faculty perceived this PEG software and to assess the degree to which they were willing to accept its use in their classes.

METHODOLOGY

Data Collection

A 36-question survey was developed to assess how faculty felt about PEG software and how willing they were to use the technology. The survey initially coded the responses in four domains; time commitments (T), value of feedback and grades (V), and ethical issues (E) (including cheating). These issues were consistent with the literature evaluating what faculty considers when evaluating the grading process. The fourth domain, technology acceptance (C) was added as a measurement since this was considered a possible causal factor in whether or not faculty would accept the use of this technology. Appendix A illustrates the survey with initial coding of factors. The survey used 26 Likert scaled statements along with 9 demographic questions and one open-ended question. The surveys were distributed at three universities; one large private university, one medium sized private university and one large public university. Each school also had large amounts of online course offerings taught by faculty. After gaining institutional approval at the subject universities, surveys were then mailed to faculty departments for distribution.

Data collection took place over a nine week period. All survey data was transferred onto an Excel spreadsheet and then uploaded in SPSS (Statistical Package for the Social Sciences, version 21) for analysis. To minimize the possibility of data entry error, the process of transferring the data was done twice by different members of the research team.

Data Analysis

One hundred and seventy surveys were returned. After the data were transferred in SPSS it was checked for errors and outliers. Frequency distributions and descriptive statistics were run, and the data was then analyzed, resulting in 5 surveys being eliminated from the data set. All were heavily water damaged and the data was illegible, making it impossible to determine the actual participant responses. That left 165 useable surveys. From those surveys, 12 had some missing data (usually one or two blank responses); however, they were retained for this study. The results from the survey are listed in Table 1:

Table 1: Survey Demographic Data

Gender		Percentage of Sample	
Male	104	63%	
Female	61	37%	
Age			
Age Range	29 to 78	Median Age	47
Academic Discipline		Percentage of Sample	
Business	61	37.0%	
Education	14	8.5%	
Computer Science	19	11.5%	
Communication	11	6.7%	
Social Sciences	18	10.9%	
Engineering	6	3.6%	
Other	36	21.8%	
Have you ever taught a class that utilized computer grading software for written work		Percentage of sample	
Yes	14	8.4%	
No	150	90.1%	
Which of the following types of courses do you primarily teach?		Percentage of sample	
On campus	111	67.2%	
Online	52	31.5%	
Have you ever taught an online course		Percentage of sample	
Yes	114	69.1%	
No	50	30.3%	
Years of College level teaching experience		Percentage of sample	
0 to 5	11	6.7%	
6 to 10	45	27.2%	
11 to 15	47	28.5%	
Over 15	48	29.1%	
Faculty Status		Percentage of sample	
Full-time	39	23.6%	
Part-time	120	72.7%	
Tenure tack status		Percentage of sample	
Tenured	21	12.7%	
On Tenure Track	14	8.5%	
Not Tenured	124	75.2%	

Table 2 illustrates the survey results for the Likert scaled statements:

Table 2: SD = Strongly Disagree, D = Disagree, A = Agree, SA = Strongly Agree

	< 24 hours (1)	1 to 7 days (2)	> 1 week (3)	> 2 weeks (4)	Mean
How long does it usually take to grade a short essay (less than 10 pages)	25	88	41	11	2.23
How long does it usually take to grade a major essay (10 pages or greater)	8	55	80	22	2.70
Statement	SD (1)	D (2)	A (3)	SA (4)	Mean
I believe that grading student written work allows me to form a greater bond with students.	20	39	81	25	2.67
Grading and evaluating student written work in a timely manner is important to me.	2	10	118	35	3.13
Technology can be used to simulate how faculty grade.	33	71	44	17	2.27
Providing substantive feedback on student written work is important to me.	8	31	110	15	2.80
I believe that students value my feedback and incorporate it into future assignments.	31	48	66	20	2.45
Grading student written work takes considerable time.	9	35	60	61	3.05
Grading of student written work is highly subjective.	16	41	88	19	2.67
I would accept the use of computer grading for student written work.	36	69	47	12	2.21
Properly designed computer programs can simulate human responses to almost any stimuli.	39	50	51	25	2.38
Effective and timely grading of student work is part of my performance evaluation.	31	109	20	5	1.51
Faculty bias in grading of student written work is a serious issue for me.	16	100	30	19	2.32
I enjoy grading student written work.	20	50	53	32	3.07
I value the personal nature of feedback on student written work.	27	44	56	38	2.67
Computer generated grades and feedback on student written work would be of equal value to that given by an instructor.	41	98	13	11	1.96
A computer program is capable of effectively evaluating and grading student written work.	38	72	30	15	2.01
I would prefer to use a computer program to grade student written work.	35	81	31	18	2.19
Computer generated grades and feedback are appropriate for online courses.	13	75	48	29	2.56
Computer generated grades and feedback are appropriate for campus-based courses.	20	98	30	17	2.27
Grading student written work is one of the most important aspects of teaching.	24	35	91	15	2.59
Computer generated grades are fair.	23	82	39	11	2.11
I believe that having computers grade and provide feedback to student written work would reduce the satisfaction I have in teaching.	21	37	79	28	2.71
Not having to grade student written work would increase my overall satisfaction with teaching.	26	29	77	33	2.70
A written assignment graded by a computer would have less bias.	14	71	73	7	2.44
Having a computer program evaluate student work would increase the instances of cheating.	9	30	35	91	3.26

The 21 Likert scaled statements (numbers 3 through 23 on the survey) were subjected to Principal Component Analysis. Although the survey had undergone face validity testing from 2 subject matter experts outside the research team, since this was the first use of the instrument, determining if the grouped variables were properly aligned was of importance. Prior to performing PCA, the suitability of the data for factor analysis was evaluated. Inspection of the correlation matrix revealed the presence of several correlations of .3 or above. The Kaiser-Meyer-Oklin value was .72, exceeding the recommended value of .6 and Bartlett's Test of Sphericity reached significance, supporting the factorability of the correlation matrix [24].

Principal Component Analysis revealed the presence of four components with eigenvalues exceeding 1, explaining 56.8%, 22.4%, 9.8%, and 7.3% of the variance respectively. A scree test and parallel analysis were also performed in order to validate that four components contributed significantly to the structure and could be used to assess faculty perceptions on PEG software. The components were; value of grading and feedback (statements, 3,6,7,14,15,21,23,24), time commitments (statements 4, 8). Ethical considerations (statements 9,12,13,19,20,22,25,26), and technology acceptance (statements 5,11,16, 17). Two other questions (10, 18) were used as measurements of the willingness to accept PEG software. The selection of the grouped variables was done after several incarnations of the survey were reviewed by content experts. As a result of PCA, two statements shifted. Statement 12, *Effective and timely grading of student work is part of my performance evaluation* and Statement 19, *Computer generated grades and feedback are appropriate for online courses* both shifted to the factor of time commitments. This was surprising and may reflect a number of concerns with grading. The time commitment to teach online is generally considered to be greater than for campus-based courses and may be responsible for the shift. Additional testing was to done to assess the reliability of the survey. With short subscales (generally less than 10 items) reliability can be established using the inter-item correlation for the items. The recommended optimal range for inter-item correlations is .2 to .4 [24]. Each of the four subscales was within those parameters. After establishing the validity and reliability of the survey instrument, the demographic variables were cross tabbed with individual statements on the survey as well as grouped factors in order to evaluate any relationships between the variables.

FINDINGS

Independent-sample t-tests were conducted to compare the four grouped perception variables with each of the demographic variables as well as the two grouped variables measuring acceptance of PEG software. Most of the demographic variables showed no significant differences. Three variables, *Have you ever utilized computer generated grading of written assignments?*, *Academic Discipline*, and *Which of the following types of courses do you primarily teach, online or on campus?* reached statistical significance ($p < .05$) and had moderate or high effect sizes (using eta squared as measurement). These variables were then subjected to ANOVA testing. A one-way between-groups analysis of variance was conducted to examine the impact of each variable and the degree to which they were willing to accept PEG software in their classes. Each reached statistical significance and the effect sizes are summarized in Table 3:

Table 3: Analysis of Effect of Variable on Acceptance of PEG Software

Variable	Eta Squared	Analysis of Relationship
Have you ever utilized computer generated grading of written assignments?	.37	High
Academic Discipline	.06	Moderate
Which of the following types of courses do you primarily teach, online or on campus?	.22	High

Overall, academic discipline did have an effect on acceptance of PEG software. A breakdown by discipline illustrates an interesting aspect of this phenomenon (Table 4):

Table 4: Cross tab of academic discipline (survey item 30) and I would accept the use of computer grading for student written work (survey item 10)

Academic Discipline	Strongly Disagree	Disagree	Agree	Strongly Agree
Business	8	25	21	6
Education	3	5	5	1
Computer Science	7	11	1	0
Communication	3	4	3	1
Social Science	6	5	5	2
Engineering	1	2	3	0
Other	8	17	9	2

While the majority of faculty were not supportive of PEG software (64% either disagreed or strongly disagreed with the statement), this differences was most pronounced among Computer Science faculty, with 18 of the 19 participants either strongly disagreeing or disagreeing with the statement. This was a surprising finding since it was assumed that computer science professors would have a natural inclination to be more accepting of this technology. Faculty that had experienced using PEG software also represented a mixed response in their acceptance of PEG software. Of the 14 faculty that had used this software, only 6 responded positively on the survey (either agreeing or strongly agreeing with statements 8 and 18). Faculty that taught primarily campus-based courses were also less accepting of the use of PEG software. 71 of the 111 (64%) campus-based faculty had a negative perception of PEG software compared with 27 of the 52 (51.9%) online faculty members. Overall, faculty were not supportive of the use of PEG software in their courses. This was also illustrated by the open-ended question (survey item 27). Only 17 people responded to this question, however, the responses were universally negative and included statements such as “this just makes it easier for them to cheat”, and “I don’t trust software to grade essays”. One of the more interesting insights was expressed in this section, “This software totally changes the relationship faculty will have with both students and the administration. Administrators already think most faculty ‘knowledge’ can be captured and used and therefore once grading is also removed from faculty control, what is left for us to do?”

DISCUSSION

The study revealed some interesting insights into the perceptions that faculty have regarding the use of PEG software. The degree to which Computer Science faculty rejected the use of this software was among the most interesting aspects of the study. While there was only a small number of such faculty in the participant group, these findings indicate that further research into why computer science experts were more overwhelmingly opposed to PEG software would prove instructive. It is possible that their level of expertise allows this demographic group greater insight into some of the potential problems associated with this technology. One of the most interesting aspects of the study was the difference in teaching platform (online versus campus-based) and acceptance of PEG software. While no demographic group was accepting of this technology, the online faculty were the closest to reaching at least 50% willing to accept it. There may be several factors that contributed to this. Time commitments were one of the grouped variables and most research indicates that online teaching requires a greater time commitment than campus-based teaching. It is possible that the desire to reduce the amount of time expended by online faculty contributed to their greater acceptance of PEG software. Also, as there is limited face to face interaction in virtual classrooms, faculty may not feel as ‘vested’ in individual students vis-à-vis campus-based experiences with students. Again, this represents an interesting area for future study. The growth of online course offerings and the potential of PEG software to reduce the time commitment for online teaching represent a strong potential market for this software and gaining greater acceptance from this faculty demographic could be a critical factor in the success of PEG software.

CONCLUSIONS

While it seems apparent that most faculty have strong perceptions against accepting PEG software into their classes, there were some demographic factors that impacted the effect size of these perceptions. Further study on these factors and increased analysis on the relationship between both student and faculty perceptions on PEG software will remain an important area of study. Certainly the software will continue to develop and improve and there may

5. Curran, M., Draus, P., & Maruschock, G. (2013). Student perceptions of project essay grade (PEG) software. *Issues in Information Systems*, 14(1), 89-98.
6. Cavanaugh, J. K. (2006). What did you get? A faculty grade comparison. *Quality Assurance in Education: An International Perspective*, 14(2), 179-186.
7. Cole, W. (1993). By rewarding mediocrity we discourage excellence. *The Chronicle of Higher Education*, B1-B2.
8. Farley, B. L. (1995). "A" is for average: The grading crisis in today's colleges. ERIC Document Reproduction Services (No. ED 384 384).
9. Gaultney, J. F. & Cann, A. (2001). Grade expectations. *Teaching of Psychology*, 28, 84-87.
10. Giles, J. (2012). Software gets top marks for grading student essays. *New Scientist*, 214(2861), 21.
11. Gordon, M. E., & Fay, C. H. (2010). The effects of grading and teaching practices on students' perceptions of grading fairness. *College Teaching*, 58(3), 93-98.
12. Greville, E. C. (2009). A rose by any other name: Grading and Assessment. *Assessment Update*, 21(5), 1-2.
13. Guskin, S. L., Peng, C. J., & Simon, N. (1992). Do educators react to "multiple intelligences"? Effects of educators' stereotypes on judgments and expectancies for students with diverse patterns of giftedness/talent. *Gifted Child Quarterly*, 36(1), 32-37.
14. Harrison, T. (2007). My professor is so unfair: Student attitudes and experiences of conflict with faculty. *Conflict Resolution Quarterly*, 24(3), 349-368.
15. Holmes, L. E., & Smith, L. J. (2003). Student evaluations of faculty grading methods. *Journal Of Education For Business*, 78(6), 318-323. doi:10.1080/08832320309598620
16. Hsu, T. C. & Shermis, M. D. (1989). The development and evaluation of a microcomputerized adaptive placement testing system for college mathematics. *Journal of Educational Computing Research*, 5(4), 473-485.
17. McLauchlan, W. (1999). Educational testing service responds. *College Teaching*, 47(1), 7-8. doi:10.1080/87567559909596069
18. Page, E. B. (1966). The imminence of grading essays by computer. *Phi Delta Kappan*, 48(1), pp. 238-243.
19. Page, E. B. (1994). Computer grading of student prose, using modern concepts and software. *Journal of Experimental Education*, 62(2), pp. 27-142.
20. Page, E. B. (1996) Grading essays by computer: Why the controversy? Paper presented at the annual meeting of the National Council on Measurement in Education, April New York, NY.
21. Page, E. B. & Petersen, N. S. (1995). The computer moves into essay grading: Updating the ancient test, *Phi Delta Kappan*, 76(7), pp. 561-565.
22. Page, E. B., Lavoie, M. J. & Keith, T. Z. (1996). Computer grading of essay traits in student writing. Paper presented at the annual meeting of the National Council on Measurement in Education, April, New York, NY.
23. Page, E. B., Poggio, J. P. & Keith, T. Z. (1997). Computer analysis of student essays: Finding trait differences in the student prose. Paper presented at the annual meeting of the American Educational Research Association, March, Chicago, IL.
24. Pallant J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS for windows* (3rd edn). New York: McGraw Hill.
25. Schnelle, R. (2009). Taking the pulse of the essay. *Writing on the edge*, 19(2), 79-84.
26. Schutz, K. (2012). *A comparison of community college full-time and adjunct faculties' perceptions of factors associated with grade inflation*. ProQuest Information and Learning Company, LLC. (UMI No. 3541624).
27. Schutz, K. R., Drake, B. M., & Lessner, J. (2013). Do community college full-time and adjunct faculties differ in their perceptions of rigor in assigning grades? *American Journal Of Business Research*, 6(2), 59-79.
28. Shermis, M. D., Mzumara, H. R., Olson, J., & Harrington, S. (2001). On-line grading of student essays: PEG goes on the world wide web. *Assessment & Evaluation In Higher Education*, 26(3), 247-259. doi:10.1080/02602930120052404
29. Shermis, M. D., Mzumara, H. R., Lillig, C. & Brown, M. (1997). Computerized adaptive testing through the World Wide Web. Paper presented at the annual meeting of the American Psychological Association, August, Chicago, IL.
30. Shermis, M. D., Rasmussen, J. L., Rajacki, D. W., Olson, J. & Marsiglio, C. (1999). Prompts and themes as sources of variance in grading college placement essays. Manuscript submitted for publication.

31. Shores, M., & Weseley, A. J. (2007). When the A is for agreement: Factors that affect educators' evaluations of student essays. *Action In Teacher Education*, 29(3), 4-11.
32. Smith, W. (1993) Assessing the reliability and adequacy of using holistic scoring of essays as a college composition placement technique, in: M. M. Williamson & B. Huot (Eds) *Validating Holistic Scoring for Writing Assessment: theoretical and empirical foundations*, 142–205 (Cresskill, NJ, Hampton Press).
33. Strong, B., Davis, M., & Hawks, V. (2004). Self-grading in large general education classes: A case study. *College Teaching*, 52(2), 52-57. doi:10.3200/CTCH.52.2.52-57
34. Viadero, D. (1995). Making the grade. *Education Week*, 14(36), 33.