Homework Graded by Students

Dr. Ashraf Badir P.E., Florida Gulf Coast University

Dr. Badir is an Associate Professor in the Environmental and Civil Engineering Department at the U.A. Whitaker College of Engineering in Florida Gulf Coast University. He earned his B.Sc. (1982) in Civil Engineering and M.Sc. (1985) in Structural Engineering from Alexandria University, Egypt. He also holds a M.Sc. (1989) and a Ph.D. (1992) in Aerospace Engineering from Georgia Institute of Technology.

Dr. Robert O'Neill P.E., Florida Gulf Coast University

Dr. ROBERT (BOB) J. O'NEILL is Professor and Chair of the Department of Environmental and Civil Engineering, U.A. Whitaker College of Engineering, Florida Gulf Coast University. He received a B.S. from the United States Military Academy in 1975, an M.S. in Structural Engineering and an M.S. in Geotechnical Engineering from Stanford University in 1984 and a Ph.D. in Structural Engineering from Kansas State University in 1993. Prior to his coming to FGCU he was a Professor of Engineering at Roger Williams University and an Associate Professor and Director of the Civil Engineering Analysis Group at the United States Military Academy. Dr. O' Neill is a retired Lieutenant Colonel, U.S. Army Corps of Engineers. He has been active at the national level with ASCE's Technical Council on Computing and Information Technology (TCCIT), Committee on Faculty Development (CFD) and Excellence in Civil Engineering Education (ExCEEd) initiative. Dr. O'Neill is a licensed Professional Engineer in California, Florida, Nevada and Virginia. He is a civil engineering program evaluator for the Accreditation Board for Engineering and Technology (ABET). He is an American Society of Civil Engineering Fellow (ASCE), a member of the American Society for Engineering Education (ASEE), and Phi Kappa Phi National Honor Society.

Homework Graded by Students

Abstract

Permitting senior level university students to grade their individual assignments has the potential for numerous benefits to both students and faculty. However, does such a practice seem right? This paper reports the results of a first trial, implementing students self-grading in a reinforced concrete design civil engineering course. The study provides a summary of students' perception concerning self-grading and lessons learned.

Introduction

Learning is a dynamic progressive process reinforced by making mistakes. At Florida Gulf Coast University, the faculty in the U.A. Whitaker College of Engineering have adopted a combined lecture lab format. The professor initiates a lesson by explaining the first part of a topic, solving a relevant problem, then letting the students work in groups solving other problems on the board while the instructor goes around answering students questions, providing guidance and/or correcting their mistakes as needed. Thereafter, the faculty continue with the succeeding part of the lesson. Throughout the years, students have positively evaluated this format emphasizing that instantaneous correction of errors by the instructor or the students' peers is of a tremendous benefit to the students and helps them understand the subsequent professor's explanation within the same lecture.

Students rarely use their graded exams to study and many just quickly go over their graded assignments without giving the necessary attention to their errors, let alone discovering alternative solution methods to the same problems. The reality is that many students seldom use the opportunity to learn from their homework/exam mistakes. The loss of learning opportunity from TA or teacher graded homework prompted the idea of letting the students discover their own mistakes through self-grading of assignments and even mid-term exams in a reinforced concrete design course taught in the civil engineering program.

Self-grading one's own educational work allows the individuals to quickly identify his or her oversights and provides the students with faster and more detailed feedback regarding their academic performance (Weaver & Cotrell, 1986). Moreover, it provides a realistic sense of their own strengths and weaknesses and that they can use knowledge of their own achievements to direct their studying into productive directions (Boud, 1986; Schön, 1983, 1987). This approach might not work in courses requiring integrative analyses, theory, syntheses, or interpretive skills (Simkin 2015). However, self-grading can be used successfully in courses that focus on problem-solving techniques, where students are given a grading rubric with which to evaluate their work (Boud, 1989; Panadero and Jönnson, 2013; Simkin 2015); especially if a unique and/or a series of number are the only correct answers for each step of the solution.

Methodology

An experiment has been adopted in which 27 students in a senior civil engineering reinforced concrete design course, taught in the fall semester 2016, were assigned the task to grade their own work. Students scanned and uploaded their assignment on CANVAS (learning management

software) by the deadline, usually the starting time of the class. At the beginning of the class period, the instructor displayed on the screen the solution of each problem. Students graded their own work based on a pre-allotted credit by the instructor for each segment of the solution. The students were encouraged to ask for clarifications regarding the solution and the grading scheme. The graded work was collected and the instructor had the opportunity to check, after the class period, the students' grading while briefly comparing the submitted written work with the previously uploaded one. The assignments for this course used either the end-of-chapter problems from "Design of Reinforced Concrete," by McCormack and Brown (2016) or custom assignments and exam questions created by the instructor. Students graded their homework and one question of the first exam. The homework and the exam question represented 10% and 9% of the overall course grade, respectively. Moreover, the instructor graded all homework and exams and the instructor's assigned grades were compared with students self-scores for the purpose of this study.

Determining if a student's grade is consistent with the teacher's assigned grade can be accomplished by different statistical methods (King and Cai, 2016). A common statistical method is to calculate Cohen's Kappa. In its simplest definition, Cohen's Kappa is a measurement of percent agreement between two raters (i.e., student and teacher) (McHugh, 2012). Sadler and Good (2006) also recommended additional methods for measuring the consistency between the students grades and the teacher grades: (1) rank and Pearson correlation, (2) *t* test comparing difference in mean grades, (3) effect sizes (difference in mean grades in units of standard deviation), and (4) chi-square statistic (comparing grade categories). In this study, the authors adopted the "*t*-Test: Paired Two Sample for Means" data analysis using Excel.

Sadler and Good's (2006) found that students tended to assign lower grades to their peers than to themselves, nevertheless only self-grading of assignments were used in this study to avoid any potential conflict or violation of the Family Educational Rights and Privacy Act (FERPA) rules and regulations. However, on February 19, 2002, the Supreme Court decided unanimously that peer grading does not violate FERPA in the case of Falvo v. Owasso Independent School District (2000). More details about the case is discussed by Sadler and Good (2006).

Findings

Table 1 shows the study results for a total of 8 assignments (assignment #4 was not self-graded by the students) and the first problem of the first exam (last column), along with selected statistics following the work presented by Simkin (2015) in a business class environment. As shown in the table, assignments were worth different total amounts – values that were set according to the amount of problems required for each assignment (row 2). Almost every student of the 27 registered in the class submitted their assignment as seen in the count row. Table 1 also displays the maximum difference in student-grader pair of scores (row 4). Thus, the "Max Difference" value of "5" for Assignment 8 was the largest difference observed between the student's grade and the instructor's grade for that homework. Similarly, the "Min Difference" was the smallest difference – i.e., the situation in which the grader awarded higher grade than the student did for his or her own assignment (row 5). The average difference between the instructor's grade and the student's grade for assignment 1 was 0.27 points, meaning that, on the

average students graded themselves about 0.27 points higher than the instructor did. The matched pairs *t*-statistics in Table 1 is the different-from-zero test typical of matched-pairs tests, i.e. the null hypothesis was that there were no differences in the means of the designated grades by the students and the instructor. The *t*-statistics value of 1.77 for assignment 1 in row 8 is less than the critical value 2.787 in row 9 indicating that we would not reject the null that there is no difference between the students self-graded scores and the instructor graded scores. Similar results are obtained by comparing each *t*-statistic value in row 8 with its corresponding critical value in row 9 including the exam problem. The *p* value (row 10) for assignment 1, $p = 0.09$, for every other assignment, and for the exam problem is greater than the set alpha value of 0.01, again indicating that the null hypothesis cannot be rejected.

Table 1: Assignment Grading statistics, using a matched-pairs test for each assignment.

1. Assignment #	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**P1Exam1**
2. Points	20	40	20	80	40	50	40	60	45
3. Count	26	27	27	27	27	26	27	25	27
4. Max Difference	3	2	1	N/A	4	2	4	5	4
5. Min Difference	0	0	-1	N/A	-1	0	0	-1	0
6. Average Difference	0.27	0.22	0.10	N/A	0.28	0.15	0.30	0.33	0.37
7. Standard. Dev. Of Differences	0.78	0.64	0.385	N/A	1.022	0.53	0.869	1.177	0.967
8. Matched pairs t-statistics*	1.77	1.80	1.0	N/A	1.41	1.44	1.77	1.47	1.99
9. *t* critical two-tail	2.787	2.778	2.778	N/A	2.778	2.787	2.778	2.797	2.778
10. *p*	0.090	0.083	0.327	N/A	0.170	0.161	0.088	0.153	0.057

*All results were statistically insignificant at an alpha level of 0.01.

Besides comparing students-instructor grading, students' perception about self-grading their homework was assessed by a formal anonymous survey as part of the Student Perception of Instructor (SPoI) questionnaire administered at the end of each semester for every course taught at Florida Gulf Coast University. Instructors are allowed to add questions to the SPoI, and four questions regarding self-grading were added as shown in Table 2 through Table 5. Unfortunately, since the survey is administered online by the university, the instructors do not have control on the number of students who choose to participate. Out of 27 students, 15 participated in the SPoI. All responses, as provided by the university, are included in the four tables. Tables 2 and 3 summarize students' responses to two open-ended questions: (1) "*what did you like best about grading your own homework*?" and (2) "*what did you like least about grading your own homework?*" Students were prone to like grading their own homework and discovering their mistakes with some concerns about taking time away from the class. Tables 4 and 5 summarize the student responses to multiple choices questions: (1) "*How do you think grading your own homework and the questions/answers discussion during the grading affected your understanding of the topics and problems compared to being graded by the TA?"* Two-thirds of

the respondents believed it increased their understanding and none believed it had reduced it; (2) *"Overall, how would you rate your experience grading your own homework?"* Seventy five percent believed it was either good or very good, and none believed it was bad or very bad.

Table 2: What did you like best about grading your own homework?

• Makes me go back over it and see what is wrong • learned my mistakes • You have an opportunity to review your work and see the correct way to solve the problem. • I can see where I make mistakes. • It gave me time to look over the solutions before the test. • It was good to see how I messed up on a problem. • I can see where I made mistake. Usually I would not go back and look over my homework when studying the course material. • I learned my mistakes better • I had the chance to grade myself fairly.

Table 3: What did you like the least about grading your own homework?

• Takes time away from class • learning my grade lol • Sometimes you do not know how to grade, but it's not a big deal. • N/A • I think it took a lot of class time. • Design problems were difficult to grade due to the number of possible solutions. • Nothing. • It wasn't bad so no complaints • A lot of work sometimes

Table 4: How do you think grading your own homework and the questions/answers discussion during the grading affected your understanding of the topics and problems compared to being graded by the TA?

Response Option	**Percent**
Much Higher	66.67%
About The Same	33.33%
Much Less	0%

Table 5: Overall, how would you rate your experience grading your own homework?

Response Option	**Percent**
Very Good	33.33%
Good	41.67%
Neutral	25%
Bad	0%
Very Bad	0%

Discussion

As seen in the previous section, there was very good agreement between the grades assigned by the instructor and the students. The difference is statistically insignificant, and would not alter the overall final course grade of any of the students. The instructor believes that similar results would have been obtained had the comparison been between the grading of the instructor and a TA. While it is premature to generalize this result, due to the limited number of samples, one may reflect on the findings to deduce some lessons learned. The sample course was a senior engineering one in which almost all the students had previously attended one or two courses with the same instructor in a relatively small university where faculty-student interaction is stronger than a larger size university with more students attending classes. The instructor gave a very precise scheme for assigning grades to every step of the solution and answered all students' questions regarding how to grade under different situations. King and Cai (2016) state that many research studies have reported a high level of agreement between the grades by teaching staff and the grades from their students, when students could understand the teacher's requirements for assignments and grading methods. In this study, the total assigned grades for the homework was only 10%, not too high to become an incentive for a student to inflate his or her grades. It was common that students would write comments like *"Oooops… How can I do such a mistake!"* or *"Be careful next time!"* and even *"Stupid mistake*!" on their own graded homework. In many instances, the students were not sure how much to penalize themselves and the instructor would guide them by questioning whether their mistake was considered, in their opinion, a crucial one, one that shows misunderstanding of the topic or a minor error. In one instance a student, who failed to position the steel in the tension zone of a cantilevered beam, was voted by his colleagues to receive a very heavy grade penalty. Actually, the students voted their colleague *"to go to jail"* since the cantilever beam without tension reinforcement would collapse. This collaborative environment benefitted the students learning, exposed them to pitfalls that must be avoided and created a friendlier class environment. In many cases, the instructor provided different ways to solve a problem to accommodate all possible solution scenarios, which allowed the students to discover alternative ways of solving the same problem and reinforcing their overall understanding of the topic. This method not only strengthened the lesson but also realized whether the students have grasped the material and are prepared to move on.

It was observed that out of 8 homework assignments assigned to the 27 students with a total of 216 to be turned in, only 4 were not submitted or 1.9%. This is an extremely low percentage compared to the rate of 6.6%, 11.7% and 16% in the preceding three fall semesters for the same course. One may presume that a student, knowing that he or she will grade his own homework, would not want to miss submitting it and has a strong incentive to complete it on time. It is very frustrating for a student to sit in a class watching his classmates accumulating points without his or her involvement. Even if a student missed completing his or her assignments, attending the in-class grading discussion could still be of benefit.

One of the issues raised in the students' survey is the concern that self-grading takes time from class, a legitimate worry, however in many instances the time was spent in reinforcing the

understanding of the students or clarifying areas that were not well understood by some. Nevertheless, one might reduce the amount of time spent in the class by allowing the students to grade some of the homework questions or all of them on their own, outside of the class, and then quickly address any raised grading issues within the class.

Unfortunately, the textbook used in this course was introduced for the first time in fall 2016 and all the exams were closed book while providing the students with a reference data card with all needed equations and copies of the code. Previously the exams were open book and as such, the assessment of student performance on exams compared to the previous semesters where students' self-grading was not used would not be satisfactory.

Summary and Conclusions

A noteworthy illustration of self-learning is one that permits students to evaluate their own homework. While there are many benefits of such a procedure, uncertainties linger about scoring inaccuracy and truthfulness in the practice. To explore these questions in an engineering setting, the authors required the students in a senior level, reinforced concrete design course, to grade in class seven of their own homework assignments and one exam question. All homework assignments were then turned in and regraded by the faculty. Using matched pair *t*-tests, the authors found that the difference in the grades were extremely small and statistically insignificant. The general inference is that students were competent in accomplishing the evaluation tasks required and that this grading scheme enhances the overall students' understanding of the material as outlined in the students' survey. This is achieved by (1) discovering their own mistakes and correcting them, (2) reducing the percentage of the students not turning in the assignment on time, (3) getting immediate feed-back in contrast to waiting for the assignment to be graded, with the possibility of not even benefitting from the grader's remarks, (4) enticing all the students to raise questions regarding the homework solution, and (5) benefitting the students by exposing them to alternative ways to solve the same problem.

One of the primary limitation to the use of self-grading in class is the use of class time to complete them, which detracts from the time available for other activities. A possible alternative would be to reduce the amount of time spent in the class by allowing the students to grade some of the homework questions or all of them on their own outside of the class and quickly address any raised grading issues within the class.

Duplicate studies are required to confirm these conclusions and identify other classes that can profit from this strategy.

References

Boud, D.J. (1986) Implementing Student Self-Assessment, Green Guide No. 5, Sydney: *Higher Education research and Development Society of Australia.*

Boud, D.J. (1989) The Role of Self-Assessment in Student Grading, *Assessment and Evaluation in Higher Education*, 14, 20-30.

Falvo v. Owasso Independent School District (2000), *Court of Appeals*, 10th. U.S. Circuit.

King, Hunter C. and Cai, Qijie (Vicky) (2016). Self-Grading: A Commentary, *iSALT Resources: Theories*, Concepts, and Measures. Paper 7. http://cornerstone.lib.mnsu.edu/isalt_resources/7

McCormac, Jack C.and Brown, Russell H (2016) *Design of Reinforced Concrete*, 10th Edition, John Wiley & Sons, Inc.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

Panadero, E. & Jönsson, A. (2013). The Use of Scoring Rubrics for Formative Assessment Purposes Revisited: A Review, *Educational Research Review*, 9, 129-144.

Sadler, P. M., & Good, E. (2006). The impact of Self-and Peer-Grading on Student Learning. *Educational Assessment*, *11*(1), 1-31.

Schön, D.A. (1983) *The Reflective Practitioner: How Professionals Think in Action*, New York: Basic Books.

Schön, D.A. (1987) *Educating the Reflective Practitioner: Towards a New Design for Teaching and Learning in the Professions*, San Francisco: Jossey-Bass Publishers, Inc.

Simkin, Mark G. (2015). Should you allow students to Grade their own homework? *Journal of Information Systems Education*, Vol. 26(2)

Weaver, R. L & Cotrell, H.W. (1986). Peer evaluation: A case study. *Innovative Higher Education, 11,* 25–39.