Should you Allow your Students to Grade their own Homework?

Mark G. Simkin
College of Business
University of Nevada
Reno, Nevada 89557, USA
MarkGSimkin@yahoo.com

ABSTRACT

Allowing students to grade their own homework promises several advantages to both students and instructors. But does such a policy make sense? This paper reports the results of an experiment in which eight separate assignments completed by approximately 80 students were first graded by the students using a grading rubric, and then re-graded by a teaching assistant, using this same rubric. The study found that the differences observed in the two sets of data were statistically significant, but (in the author's opinion) acceptably small. The study also confirmed observations by earlier researchers that students who generously grade their work tend to fall among the lower-performing individuals in a class.

Keywords: Collaborative learning, Self-grading, Assessment, Self-assessment, Homework grading

1. INTRODUCTION

The term *collaborative learning* (CL) refers to instructional venues in which students assume responsibility for some of the educational activities in their courses. Although such collaboration often occurs among students—as for example, in group discussions and decision making—CL can also involve collaboration between students and their professors. For example, Vander Schee (2011) suggests that allowing students to select the weights used to determine their final course grades positively influences their commitment to their courses, increases their sense of control of their classes, and even improves course performance.

A growing body of both anecdotal and empirical evidence suggests that collaborative learning applications enjoy many advantages. For example, both Koppenhaver (2006) and Taneja (2014) argue that tasks requiring teamwork increase opportunities for collaboration, help students learn how to reach collective decisions, improve interpersonal skills, and facilitate group problem solving. Similarly, Iqbal, Kousar, and Rahman (2011) note that collaborative learning exercises may be effective strategies in distance learning environments, in which face-to-face interactions are limited.

Of particular interest to this author is the idea of allowing students to grade their own homework. This approach might not work in courses requiring integrative analyses, theory syntheses, or interpretive skills. But a growing body of empirical evidence suggests that self-grading can be used successfully in courses that focus on problem-solving techniques and where students are given a grading rubric with which to evaluate their work (Boud, 1989; Panadero and Jonnson, 2013). Similarly, student grading may become more feasible where homework problems have strict, right answers such as in accounting, IS, or the STEM disciplines (science, technology, engineering, and mathematics).

"Self-grading" promises advantages to both students and instructors. One potential benefit to *students* is its ability to increase their engagement and commitment to the learning goals of a course. Self-grading also provides immediate feedback—a benefit that can positively influence learning and increase retention (Edwards, 2007). Student self-grading also provides an opportunity for students to deepen their understanding about a subject—for example, to better understand why a given answer is wrong, or why an alternate answer is better (Sadler and Good, 2006; Cherepinsky, 2011).

A growing body of empirical evidence also suggests that self-grading improves class attendance, makes the classroom experience a friendlier, more productive, and cooperative environment, reduces student-teacher conflict, decreases student anxiety, and provides a shared sense of ownership for the learning process (Strong et al., 2004, Edwards, 2007). Studies also suggest that student self-assessment has the potential to transform a student's view of education from a passive, external experience to an internalized value of lifelong self-learning (Dungan and Mundhenk, 2006). Finally, studies suggest that self-grading can enhance student self-esteem and confidence, motivate them to learn, and increase positive attitudes about a course and the instructor who teaches it (McVarish and Solloway, 2002; Strong et al., 2004).

One potential benefit of student self-grading to *instructors* is the ability to assign homework that the professors might otherwise not require—a characteristic of special advantage to teachers of large classes and a policy that authorities list among the seven best practices of teaching (Chickering and Gamson, 1987; Geide-Stevenson, 2009). A second advantage is the time that instructors save because their work is limited to recording tasks instead of grading tasks (Sadler and Good, 2006). A third advantage is the potential to increase student engagement and transform students from passive listeners to active evaluators and motivated learners (Stefani, 1994; Mahlberg, 2015).

A fourth benefit is the usefulness of self-grading in online education, where the lack of grading resources limits what can be done by a single instructor (Ohland et al., 2012). For example, Udacity is a major MOOC provider that includes "self-grading" among its assessments (Boyde, 2013). A fifth benefit is self-grading's potential to increase student engagement in coursework, if only because self-grading transforms students from passive submitters of work to active evaluators of such work. In the author's experience, another advantage is that it enables instructors to discuss novel or creative solutions that online software might grade as "incorrect," but that an instructor can acknowledge and allow as correct in class.

Finally, self- or peer grading has the ability to perform assessment tasks that instructors cannot. For example, a professor cannot independently assess the amount of effort expended by the members of a team working on an outside project, but might want to lower the final scores of those who ride the coattails of others. This is particularly useful for the teamwork characteristic of group assignments in project management, systems analysis, or computer programming classes (Tu and Lu, 2005; Hadar et al., 2008).

Just because instructors *can* allow students to grade their own homework does not mean that instructors *should* adopt such a policy. One problem is that self-grading takes valuable class time and therefore imposes an opportunity cost. Another concern is the view that homework should be optional because it is simply a means to an end - the mastery of course materials - and that in-class tests adequately motivate students to learn them (Geide-Stevenson, 2009). Finally, some students balk at grading chores that they consider "busy work" or "not my job."

Several additional factors also negatively influence the advisability of self-grading. One of the most onerous concerns is the amount of expertise required in the grading process itself. How can students adequately grade their own work in those courses covering unfamiliar material? Several authors suggest that they cannot - see, for example, Andrade and Du (2007) and Kirby and Downs (2007). But a growing body of empirical evidence suggests the opposite. For example, a study by Boud and Falchikov (1989) found that most student marks agreed with those of their teachers. Similarly, a study by Stefani (1994) found that student self-assessed grades were similar to those of their tutors. Finally, Leach (2012) found no statistical difference between the mean student (self-assessed) grade of 5.57 and the mean teacher grade of 5.58 ($p<.01$) for the homework materials of 120 students in her adult education classes.

Lastly, there is the matter of "honesty" in student self-grading. Even if students are capable of evaluating their work objectively does not guarantee that they will do so. Moreover, if instructors include self-graded homework as a component of their final course grades, there is an obvious incentive for students to be generous in grading themselves (Andrade and Du, 2007; Kirby and Downs, 2007; Long, 2003). This matter is particularly interesting to instructors in colleges of business, where "cheating" is variously described as "common," "pervasive," or "pandemic" (McCabe, 2005; Bing et al., 2012). This concern explains why some experts believe that self-grading is inappropriate in higher education (Kirby and Downs, 2007; Thompson et al., 2005).

A growing body of empirical investigations appears to confirm this belief. For example, Sadler and Good (2006) compared student homework evaluations with teacher grades for the same work in four of their general-science classes and found that "lower-performing students tended to inflate their own low scores." Similarly, a study by Leach (2012) of 472 students made these same observations for lower-achieving students, but also found that higher-achieving students tended to underrate themselves. Yet a third study by Strong, Davis, and Hawks (2004) of 480 students in their history classes found that 57 percent of self-assessments resulted in "A" grades, compared to 31 percent of "A" grades when teachers assigned grades.

The empirical evidence on self-grading is inconsistent, however. For example, a study by Lopez-Pastor et al. (2012) involving 183 students found high correlations among self-assessed grades and professional grades in all three of the study classes. Similarly, when using grading rubrics, the study by Sadler and Good (2006) cited above found "very high correlations" (with r values between .91 and .94) between students and teachers on sets of test questions.

The claim that students can accurately and fairly grade their own assignments is a testable hypothesis, but investigations of this in business schools in general and the field of IS in particular are notable for their absence. This paper describes an empirical investigation by the author to address this question, using data from two sections of an information systems course. The next section of this paper describes this study, and the section after that discusses the study's findings. A number of concerns limit these findings, which this paper discusses in yet a further section. The final section of the paper provides a summary and conclusions.

2. A NEW STUDY

The objective of this study was to test the hypothesis (H_0) that students can accurately grade their own homework. To do so, the author conducted the following experiment in two sections of an information systems class he taught within the college of business administration of a public university in the western United States. Both classes were for the same course: a sophomore-level class in advanced Excel and Access. This course is required of IS and accounting majors, but is optional for all others.

2.1 Methodology

The homework for this course used either the end-of-chapter problems from Parsons et al. (2011) or custom assignments developed by the author, and the homework counted for 30% of the final course grade. Homework was due almost every week. For this experiment, the author chose eight Excel assignments with which to test the hypothesis that students can accurately grade their own work. The rationale for selecting these Excel tasks is that Excel spreadsheets tend to be more visual than Access, and students tend to like them more than Access exercises. The numerical outputs are also perhaps easier to grade.

Students were required to document their work on paper and homework was due at the start of each class. At that time, the correct answers for the assignment were displayed onscreen in the front of the classroom, along with the grading rubric provided previously with the homework. The author also asked students to print these rubrics as part of the cover sheets for their homework. Appendix A contains a typical assignment and an exemplary grading rubric.

Students graded their own work, and were free to award themselves partial credit if they thought they deserved it. The instructor then collected the assignments and asked his (very competent) teaching assistant (TA) to grade them again using the same solution presented in class. In a few cases, the TA gave students *more* points than what they had awarded themselves. In others, he gave fewer points.

2.2 Findings

Table 1 shows the study results, along with selected statistics. As shown in the table, assignments were worth different total amounts - values that the author set according to the difficulty and amount of work required for each assignment. This is why the first assignment - a warm-up exercise - was only worth 20 points, while the last assignment - a comprehensive consolidation exercise - was worth 75 points.

Although there was a combined total of almost 80 students in the two sections of this class, not every student completed every assignment (this despite a weight of 30 percent towards the final course grade). Also deleted from the sample were homework grades for students who completed an assignment but were unable to come to class to grade it. These are the two primary reasons why the number ("count") of students completing each homework differed from assignment to assignment.

Table 1 also displays the maximum difference in student-grader pairs of scores. Thus, the "Max Difference" value of "7" for Assignment 1 was the largest difference observed between the student's grade and the teaching assistant's grade for that homework. Similarly, the "Min Difference" was the smallest difference - i.e., the situation in which the grader awarded the most additional points for an assignment than did the student for an assignment. The average difference between the teaching-assistant's grade and the student's grade for Assignment 1 was exactly one point, meaning that, on average, students graded themselves about one point higher than the grader did.

Finally, the t-statistic in the last row of Table 1 is the different-from-zero test typical of matched-pairs tests. Here, all values were statistically significant ($p<.01$), meaning that student grades statistically differed from (and were higher than) those of the grader. The paper discusses the materiality of these differences in the Discussion portion of the paper.

The integrative assignment given in Assignment 8 is notable for several reasons. For one, it was worth considerably more points (75) than the others. This is probably why more students completed this assignment than any other. Surprisingly, however it was *not* the assignment displaying the largest difference in student-grader scores, a dubious honor belonging to Assignment 7. Finally, the statistics for this assignment are noteworthy because it had the smallest average difference—a result that seems counter-intuitive given the maximum number of points involved.

The average difference in student and grader scores for the entire sample was ".99" - i.e., about one point per assignment. Over the eight assignments then, and using a different-from-zero test, this difference was statistically significant ($t = 5.60$, $p=.001$), meaning that the observed disparities were meaningful and unlikely to be attributable to chance. Again, the paper provides additional comments on this result in the following Discussion section of this paper.

It is also useful to determine whether only certain students had difficulty in these grading exercises. To answer this question, and for each assignment, the author identified the top 5% of students with this problem, as measured by the magnitude of the difference in their homework scores compared to those of the grader. Those students who repeatedly fell into this grouping were of special interest inasmuch as the probability of such repetition by chance is .0025 for a "double appearance" and .000125 for a "triple appearance."

Assignment:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Points:	20	25	20	20	30	30	25	75
Max Difference:	7	6	7	15	9	8	17	5
Min Difference:	-2	0	-1.5	-1	-7	0	-2	-1
Count:	34	67	64	66	66	64	61	72
Average Difference:	1.00	1.42	0.45	1.78	0.94	1.33	0.69	0.32
Std. Dev.	1.72	1.68	1.38	2.90	2.40	2.02	2.34	1.10
Sample Mean S.D.	0.30	0.21	0.17	0.36	0.30	0.25	0.30	0.13
Matched pairs t-statistic*	3.38	6.91	2.62	5.00	3.17	5.29	2.30	2.47

Table 1: Selected grading statistics, using a matched-pairs test for each assignment.
***All t-statistics were statistically significant at an alpha level of .01.**

For the sample at hand of approximately 80 students, a total of four students each appeared twice on these lists, and one student appeared three times. The final course grades for these five students were F, C, C, C, and C-. Ranking students by their final course weighted averages from 1 (top student) to 80 (lowest-achieving student), these students had class ranks of 74, 60, 64, 51, and 69, respectively. The average rank of these five students was "65" - well below the median rank of "40." These results confirm conclusions by Sadler and Good (2006) and Leach (2012) that the students who generously grade their work tend to fall among the lower-performing students in a class.

Finally, it is interesting to examine those students who awarded themselves fewer points than the grader thought they deserved. The "Min Difference" row in Table 1 reports that this happened for six of the eight assignments, and that the largest disparity was "-7" - i.e., an instance in which a student penalized himself seven points more than did the grader. There were a total of ten such occurrences for the experiment as a whole. In this group, only one student penalized his work more than once. The grades for these (nine) students were: C, B, C-, F, B+, B+, B-, B, and B+ (for the student who under-graded himself twice). The class ranks for these nine students were 66, 45, 68, 75, 27, 21, 52, 23, and 39 respectively, and the average class rank was "46." None of these students received a final, course grade of A, despite the fact that the author gave 13 such letter grades to the students in these classes. Thus, and in contrast to Sadler and Good (2006) or Leach (2012), both of whom found that higher-achieving students tended to underrate themselves, the course-grade profile of these particular students suggests that these questionable downgrade penalties were more likely attributable to simple grading mistakes than to systemic errors.

3. DISCUSSION

The most important question to answer for this experiment is how to interpret the results. The fact that the pairs of grades on all homework sets were statistically different from each other suggests that the students in these classes were unable to accurately grade themselves, or perhaps were unwilling to do so. However, in the author's opinion, several considerations mitigate such inferences.

One additional factor to consider is that, across the entire sample of paired grades, the average difference between student grades and teaching-assistant grades was about one point—a differential of five percent for 20-point homework assignments and less than five percent for homework worth more than this amount. This one point seems small, and in auditing terms, almost immaterial. If most of the students had wanted to be generous with themselves, for example, it seems more likely that this average differential would be larger - perhaps 30 percent or more. The fact that it was small is notable.

If a solution was not completely correct, both the students in this class and the grader had the latitude to award partial credit. Thus, another factor to consider is the potential variability in how such partial credit might be taken. To illustrate, consider a simple payroll problem in which employees earn simple compensation equal to pay rate times hours worked for all work less than or equal to 40 hours, time and a half for all work between 40 and 60 hours, and double time for all work over 60 hours each week. In Excel, these requirements are easily expressed in a single, nested IF formula.

Now suppose that a student who works on this task constructs a formula that only computes the correct pay for the first of these three possibilities. Strict constructionists (and probably all the fictitious employees who worked overtime and got shortchanged in this problem) would argue that the resultant formula is dead wrong and deserves no credit. But most instructors would probably award partial credit for this work - for example, one point out of three - on the grounds that the student "had the right idea" or that "the working formula correctly computed gross pay one third of the time." Finally, a student might reason "I worked on this formula for an hour and deserve some credit for my efforts."

The author often received queries about such matters in class, suggested partial credit amounts for such problems, and encouraged students to grade themselves objectively. The small average grading differential observed here suggests that they tried to do so. However, not every student asked about this in class, and the lack of guidance for some mistakes probably added to the variability in the grading efforts. In the opinion of the author, a one-point differential appears more reasonable under such circumstances.

One additional factor may account for the grade differences found in this study - the potential grading variability inherent in any task requiring subjective judgment. The grader in this study used a grading rubric and a comprehensive solution key for all assignments, a strategy that the author hoped would limit assessment inconsistencies in this experiment. The fact that, in this semester, no student complained about his or her adjusted score lead the author to believe that the grader was fair and impartial in performing his work. Nonetheless, the potential for grading variations is still possible - a confound that again can increase student-grader scoring differentials.

Other factors that may also explain grading disparities in this experiment include the varying nature of underlying course materials, the relative maturity of the student sample, potential differences in the difficulty of the assignments, the grading stringency of the evaluators, the number of grading components in an assignment, and (sadly) perhaps the teaching capabilities of the author. The author believes that all of these factors potentially confound the generalizability of these findings. More research would be useful in alternate venues.

Finally, the author recognizes that the results found here differ from those found by Stefani (1994) and Leach (2012), both of whom found no statistically-significant differences between teacher and student homework scores in similar experiments using student-and-grader data. One possible explanation for this is that both prior researchers used simple difference-of-means tests to reach their conclusions. The author achieved a similar, "no-difference" result for most of the assignments using such a test when he first analyzed his data. The problem with simple difference-of-means tests is that they violate the requirement that the samples are independent. In fact, they are not. Instead, they are *pairs of observations* drawn from the same set of papers (i.e., the same population). But this information is lost with simple difference-of-means tests. For this reason, he believes that matched-pair tests are the better experimental design for this type of investigation.

4. LIMITATIONS

Several factors limit the findings presented here. Perhaps the most important is the fact that not all university classes naturally lend themselves to self-grading - for example, seminar courses or courses requiring higher cognitive processes. In such settings, self-grading homework is likely to be problematic and perhaps ill advised. However, many of the courses in the business curriculum, including IS classes and the STEM disciplines, may be more appropriate settings for self-graded work. More work needs to be done to identify the settings or subject domains in which self-grading makes sense.

Within the confines of the experiment discussed here, another concern is the fact that the findings were from only two classes in one subject and at most 80 students, taught by the same instructor during one semester and at one university. The fact that the homework in total counted 30 percent of the final grade, and not a higher percentage, may also have affected student behavior. For example, a higher percentage might have provided stronger incentive for lenient self-grading.

Another concern is that the experiment reported here involved Excel problems and therefore software that often self-corrects many of the syntax and grammar errors commonly committed, and perhaps not auto-corrected in alternate IS domains or business courses. This is a concern because student errors might intensify in such cases and therefore require greater grading expertise than that required in this experimental setting.

A third concern is the natural variability inherent in any subjective grading task. For example, the students in these classes could award themselves partial credit for work that was partially correct. Inasmuch as this variability would likely differ from student to student, so would the points they might award for the same incorrect answer. The grading rubrics and in-class discussions of both the right answers and common errors attempted to impose standard penalties for mistakes. But they are unlikely to have completely controlled for inconsistencies in grading. This might explain the statistical significance of the TA-student grading differentials.

Finally, the author returned the re-graded homework to students in each class following the period in which it was initially submitted and students were therefore able to see what "final grade" they had received for each assignment. But this policy potentially introduces demand effects in the experimental design - i.e., the likelihood that students would adjust their grading based on feedback from earlier homework assessments. (A similar concern might also apply to the grader, who might also have adjusted his grading leniency over time.)

The author hoped that grader feedback would encourage students to be more careful in future self-assessments, but the statistics provided in Table 1 suggest that such hopes were optimistic. The average differential between self-generated grades and TA grades did not seem to be much affected by such feedback, and (as noted above) continued to average about one point for each of the first six assignments. Similarly, in all 8 weeks of the experiment, the grading differential continued to be statistically significant - a disappointing result to those hoping for closer scores.

5. SUMMARY AND CONCLUSIONS

An interesting example of collaborative learning is one that allows students to grade their own homework. Although there are many advantages of such a policy, questions remain about grading accuracy and honesty in the process. To investigate these questions in a business setting, the author required the students in two sections of an advanced IS class in Access and Excel to grade eight of their own homework assignments, using a grading rubric and after viewing the correct answers. All homework assignments were then collected and re-graded by a teaching assistant using the same grading rubric and answers.

Using matched pair t-tests, the author found that, on average, student grades exceeded those of the teaching assistant by about one point - a disparity that was statistically significant but amounted to a grading differential of five percent or less. Given how small this disparity was and also after considering how many ways this variance could happen (e.g., student confusion, difficulty of the assignment, toughness of the grader, etc.), the overall conclusion is that there was little evidence to suggest that students were incapable of performing the evaluation tasks required of them, or that they were necessarily dishonest in their assessments.

Within the confines of this experiment, the overall conclusion is that students can be trusted to grade their own homework. Given the other advantages of collaborative learning afforded by such a strategy and the resultant savings in instructor grading time, in-class homework grading seems to be a "win-win" for both students and instructors. Repeated investigations are needed both to confirm this conclusion and to identify which classes can and cannot benefit from such a policy.

6. ACKNOWLEDGEMENT

The author is indebted to his teaching assistant, Mr. Ali Asudegi, for his hard work re-grading the homework involved in this experiment.

7. REFERENCES

Andrade, H. & Du, Y. (2007). Student Responses to Criteria-Referenced Self-Assessment. *Assessment & Evaluation in Higher Education,* 32(2), 159-181.

Bing, M. N., Davison, K. H., Vitell, S. J., Ammeter, A. P., Garner, B. L., & Novicevic, M. (2012). An Experimental Investigation of an Interactive Model of Academic Cheating among Business School Students. *Academy of Management Learning & Education,* 11(1), 28-48.

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

Boud, D. & Falchikov, N. (1989). Quantitative Studies of Student Self-Assessment in Higher Education: A Critical Analysis of Findings. *Higher Education*, 18(5), 529-549.

Boyde, E. (2013). Massive Open Online Courses: Time and a Little Money are a Worthy Investment, *Financial Times,* Retrieved June 5, 2015, from http://www.ft.com/cms/s/2/c101d4cc-800f-11e2-96ba-00144feabdc0.html#axzz2Y6TGqKVE

Cherepinsky, V. (2011). Self-Reflective Grading: Getting Students to Learn from their Mistakes. *Primus: Problems, Resources & Issues in Mathematics Undergraduate Studies,* 21(3), 294-301.

Chickering, A. W. & Gamson, Z. F. (1987). Seven Principles for Good Practice in Undergraduate Education. *AAHE Bulletin,* 39(7), 3-7.

Dungan, A. T. & Mundhenk, L. G. (2006). Student Self-Assessment: A Tool for Engaging Management Students in their Learning. *Teaching and Learning,* 3(1), 54-73.

Edwards, N. M. (2007). Student Delf-Grading in Social Statistics. *College Teaching,* 55(2), 72-76.

Geide-Stevenon, D. (2009). Does Collecting and Grading Homework Assignments Impact Student Achievement in an Introductory Economics Course? *Journal of Economics and Economic Education Research,* 10(3), 3-14.

Hadar, I., Sherman, S., & Hazzan, O. (2008). Learning Human Aspects of Collaborative Software Development. *Journal of Information Systems Education,* 19(3), 311-319.

Iqbal, M. J., Kousar, N., & Rahman, F. (2011). Collaborative Learning Strategies: Potential Application in Distance Education. *International Journal of Business and Social Science,* 2(12), 253-258.

Kirby, N. & Downs, C. (2007). Self-Assessment and the Disadvantaged Student: Potential for Encouraging Self-Regulated Learning. *Assessment & Evaluation in Higher Education*, 32(4), 475-494.

Koppenhaver, G. D. (2006). Absent and Accounted for: Absenteeism and Cooperative Learning. *Decision Sciences Journal of Innovative Education,* 4(1), 29-49.

Leach, L. (2012). Optional Self-Assessment: Some Tensions and Dilemmas. *Assessment & Evaluation in Higher Education,* 37(2), 137-147.

Long, R. (2003). Only an A-? Combating Grade Inflation. *Public Relations Tactics*, 10(8), 18.

Lopez-Pastor, V. M., Fernandex-Balboa, J., Pastor, L., & Aranda, A. F. (2012). Students' Self-Grading, Professor's Grading and Negotiated Final Grading at Three University Programmes: Analysis of Reliability and Grade Difference Ranges and Tendencies. *Assessment & Evaluation in Higher Education*, 37(4), 453-464.

Mahlberg, J. (2015). Formative Self-Assessment College Classes Improves Self-Regulation and Retention in First/Second Year Community College Students. *Community College Journal of Research & Practice*, 39(8), 772-783.

McCabe, D. L. (2005). Academic Dishonesty & Educational Opportunity. *Liberal Education*. 91(3), 26-31.

McVarish, J. & Solloway, S. (2002). Self-Evaluation: Creating a Classroom without Unhealthy Competitiveness. *Educational Forum*, 66(3), 253-260.

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., Layton, R. A., Pomeranz, H. R., & Schmucker, D. G. (2012). The Comprehensive Assessment of Team Member Effectiveness: Development of a Behaviorally Anchored Rating Scale for Self- and Peer Evaluation. *Academy of Management Learning & Education*, 11(4), 609-630.

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

Parsons, J., Oja, D., Ageloff, R., & Carey, P. (2011). *Microsoft Excel 2010*. Boston, Massachusetts: Course Technology.

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

Stefani, L. A. J. (1994). Peer, Self and Tutor Assessment: Relative Reliabilities. *Studies in Higher Education*, 19(1), 69-75.

Strong, B., Davis, M., & Hawks, V. (2004). Self-Grading in Large General Education Classes. *College Teaching,* 52(2), 52-57.

Teneja, A. (2014). Enhancing Student Engagement: A Group Case Study Approach. *Journal of Information Systems Education*, 25(3), 181-187.

Thompson, G., Pilgrim, A., & Oliver, K. (2005). Self-Assessment and Reflective Learning for First-Year Geography Students: A Simple Guide or Simply Misguided? *Journal of Geography in Higher Education,* 29(3), 403-420.

Tu, Y. & Lu. M. (2005). Peer- and Self-Assessment to Reveal the Ranking of each Individual's Contribution to a Group Project. *Journal of Information Systems Education,* 16(2), 197-205.

Vander Schee, B. A. (2011). Let them Decide: Student Performance and Self-Selection of Weights Distribution. *Journal of Education for Business,* 86(6), 13-49.

AUTHOR BIOGRAPHY

Mark G. Simkin is a Professor in the Information Systems Department of the University of Nevada, Reno. He earned his B. A. degree in mathematics from Brandeis University, and his MBA and Ph.D. degrees from the University of California, Berkeley. His research in end-user computing, ethics, computer education, and computer crime appears in over 100 academic journal articles, including *Decision Sciences, The Decision Sciences Journal of Innovative Education, The Journal of Accountancy, Communications of the ACM, The Journal of Business Ethics, and Communications of the Association for Information Systems.* Professor Simkin is also the author of 15 books, the most recent being *Core Concepts of Accounting Information Systems* (New York: John Wiley and Sons, 2015) with co-authors Carolyn Norman Strand and Jacob M. Rose.

APPENDIX A: AN EXAMPLE OF AN EXCEL HOMEWORK ASSIGNMENT

Objective: The purpose of this assignment is to practice using VLookup functions in Excel. The application also requires you to understand how to create spreadsheets in good form.

Description: The officers of the XYZ Company are debating what type of commission plan to use for its sales force. All suggested plans are based on the salesperson's total yearly sales.

Plan A: Plan A commissions depend upon the employee's seniority. The company uses a code from "1" to "5" to indicate seniority, with code 1 designating new salespeople and code 5 indicating the most senior salespeople. Plan A suggests the following commission schedule:

Code:	1	2	3	4	5
Plan A Commission Rate:	10%	11%	12%	15%	17%

Plan B: Some members of the executive committee do not like Plan A. They suggest the commission rate should be based on a sliding scale, with higher percentage commissions for those employees with higher total sales amounts. Plan B uses the following commission rates:

Sales:	Less than 10,000	10,000 to 29,999	30,000 to 49,999	50,000 to 74,999	75,000 or more
Plan B Commission Rate:	10%	11%	12%	15%	17%

Plan C: The president of the company suggests a compromise plan that uses both criteria. This is Plan C, with suggested commission rates as shown below. What are total commissions now? Create a separate spreadsheet to answer this question.

Deliverables: The president asks you, the company analyst, to help him decide which plan is best. Using the test data provided by your instructor, create a spreadsheet model that computes the total commissions for each plan. Which plan is the most expensive? Which plan is the least expensive? Which plan would you recommend? (Hint: The recommendation is up to you: there is nothing wrong with picking the plan that gives salespeople the most, nor is there anything wrong with saving the company money. But you must defend your recommendation in cogent writing.)

For all three problems, perform all your calculations in one or more spreadsheets that include your name, the course title, the assignment number, the spreadsheet model, and the answers to the questions above. (Hint: the tables here are NOT in precisely the format you'll need to perform the required work.) Hand in both a copy of your spreadsheets and the formulas in it. Be sure your formulas show completely in their cells.

Plan C Commission Rates

Sales↓ Code→	1	2	3	4	5
Less than 10,000	10%	10%	10%	11%	11%
10,000 - 29,999	11%	11%	12%	12%	12%
30,000 - 49,999	11%	11%	12%	12%	13%
50,000 - 74,999	13%	13%	14%	14%	15%
75,000 or more	15%	15%	16%	17%	17%

Grading Rubric for this Assignment

Item	**Maximum**	**My Score**
Plan A computations	3	
Plan B computations	3	
Plan C computations	6	
Three correct totals	2	
Identify most and least expensive option	2	
Your recommendation	1	
Formula page	2	
Totals:	**20**	

STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the Journal of Information Systems Education have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

ISSN 1055-3096