

RINGERS IN ONLINE MIS COURSES

Todd Schultz, PhD, Augusta State University, tschultz@aug.edu
James Grayson, PhD, Augusta State University, jgrayson@aug.edu

ABSTRACT

Concerns about the potential ‘ringer’ phenomenon (where a student’s online work is done by someone else) are raised frequently about courses deployed wholly or partially online. This study of both fully online and hybrid classes leads to the conclusion that if ringers exists at all, the phenomenon is slight and more prevalent for hybrid classes. While the study is too small for broad generalization, courses with a significant grade component derived from in person, validated assessments did not appear to suffer (much) from ringers doing the distance, online work.

Keywords: Online Collaboration, Academic Integrity, Online Learning

INTRODUCTION

Success with moving courses and course components online has been repeated in many locations and for various subjects. Standard development models like those outlined [3] and [6] establish the concept as accepted practice, though the successes of online instruction demonstrate the compelling nature of online instruction. Over the last six years we have been participate in this trend, overcoming – as most such efforts must – both technical and administrative hurdles.

Our paper is motivated by some feedback we received when we went though the proposal process for the new online courses. Consistently we were asked: *How do you know the student who did the work is the student who gets the credit?*

Technology enhanced cheating gets a great deal of attention, as it should. Downloading essays, copying and pasting passages without attribution, using cell phones during in-class exams are just a few examples. Our questions were more focused on the issue of validating who actually performed the work when it is clear the nature of the assignments require someone to do new, individual work.

Validation is a chronic question for online course developers, especially at institutions new to the idea where there may be administrative reluctance to the idea. Lack of a clear answer tends to weaken

perceptions of online instruction’s effect (e.g., see [1] and [6]). In our particular case, online courses require physical attendance at one or more exams taken in a computer lab where a picture ID is validated before a student takes an exam. By spreading exam periods and limiting the number of concurrent students, we get more than reasonable assurance that we are giving credit where it is due. Others use equal or stronger methods and a significant portion of the online instruction literature deals directly with authentication and validation methods [8] or student perceptions of opportunities [5].

The question lingers, however. With individuated assignments (including systems development projects and document preparation) and randomized quizzes, we are pretty sure that *someone* has done unique, creative work but is that someone the one who logged in to upload the results or take the test? We have coined the term ‘ringer’ (our use comes from Webster: *one that enters a competition under false representations*) for cases where a student obtaining the credential (e.g., earning the course or assignment grade) was in whole or part aided by someone else who did the actual work in the distance learning, on-line components.

In some ways, the issue may be larger in traditional face-to-face courses. We have never validated a student’s credentials before recording attendance, accepting assignments, or giving pencil and paper tests in traditionally taught courses, nor have most professors we know. Traditional courses provide significant hurdles to potential ringers, though: class participation is time consuming, there is the direct student-instructor interaction that goes along with exams and completing assignments which adds risk and reluctance to the prospect, and quite often the away-from-class assignments have little weight in the overall grade. In short, the cost of a ringer is higher in a face-to-face class. For on-line assignments taken at a distance, though, the hurdles are less. A friend with strong spreadsheet skills who might provide a half-hour of expert help to a student in a dorm room situation but would never consider attending a class for that same student and participating in a discussion.

Despite the quantity and quality of academic integrity related publications, we have found little established research that addresses the extent of the ringer issues. Much of the existing literature is focused on issues of plagiarism and cheating during exams which is not our central concern. We actually encourage students to build community among other distance learners (e.g., [2], [6]), especially in hybrid courses. By comparing notes and sharing approaches with others, we hope that students will learn developer skills that will allow them to succeed in their careers. This is the where the *learning* should take place, not during the validated assessments. So *collaboration* is encouraged in distance assignments but not *outsourcing* (though one might argue that an effective outsourcer should get the high-paying management job).

Our goal, then, was to find the extent to which students were using the opportunities for collaborative learning as a way to allow ringers to complete their distance, online work. In controls parlance, we were seeking *detective controls* added to the *preventative controls* already in place. By comparing how students do on similar (even identical) assignments performed or delivered anonymously online (other than login credentials) and then repeated later in class, we were able to refute the existence of ringers (at least to an extent). Our thesis or position is that by and large students commit to completing their distance work on their own and do not use “pinch hitters” when they are at a distance from strong authentication.

We describe our study and its results, finishing with some conclusions from the effort.

STRUCTURE OF STUDY

We assessed the extent of ringer phenomenon in two different venues: hybrid classes (approximately half online, half face-to-face) and online classes. In hybrid classes, much of the content comes from traditional face-to-face lectures and demonstrations but these may be supported by online modules provided as a reference. In online classes, the entire content is delivered in the form of interactive modules which simulate the software package or environment the student is learning about. Obviously, online courses work best for tools oriented courses which tend to be more introductory and hybrid methods are best for concepts rich courses which tend to be upper division.

Data for the study comes from feedback we have built into the automatic, web-based tracking systems

for our courses. We can get information on when students began and finished a component, deriving how long students had a component open. We use the frequency and durations of these encounters as measures of online effort, though it is true that in some cases students may not have been active for the entire period measured.

Figure 1 below diagrams the model we have in mind. We are not proposing here anything like a structural equations model; we simply point out some obvious interrelationships in students’ efforts and outcomes and then apply them in our data analysis.

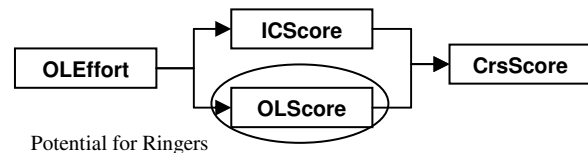


Figure 1. Sketch of Learning Model

We use each student’s course score or CrsScore (scaled to 100 pts) as a key variable which we simply associate with overall success. This is a weighted average of the online scores (or OLScore) and in class (or ICScore) scores identified below plus other assessments outside the scope of our study.

For both types of classes, portions of the eventual course grade come from assessments performed or turned in (i.e., uploaded) online. Those online submissions which had a corresponding ICScore (see below) were averaged, scaled to 100 pts and make up our OLScore variable; these are only online assignments where a similar, validated assignment was done in class. For online score we validate login credentials when these exams are taken or deliverables made but there are no authentications of exactly who is present; this potential for ringers in OLScore is the key to the issue.

We also track variables we lump together as Online Effort or OLEffort (e.g., Number of visits to modules, time spent in modules, timeliness of assignment delivery). This was intended to measure a student’s online attendance and commitment to distance work or “going to class.” We assume that OLEffort is an area where a ringer would not be asked to help; ringers would come in only for the graded OLScore assignments.

These three (CrsScore, OLEffort, and OLScore) are fairly routine measures. Our innovation comes from having in-class assessments which closely resembled or were identical (to a randomization) of work the

student had done online. Work performed in class was summarized on a 100 pt scale as ICScore. Our validation methods would prevent any ringer involvement here.

Our key hypotheses were repeated for each course type:

HA: $ICScore - OLScore = 0$ vs $ICScore - OLScore < 0$

This was a simple attempt to identify whether the ringer phenomenon exists; rejection of this hypothesis lends evidence that it might (though the difference may indeed come from other sources), since students were doing worse on the same thing.

HB1: *The correlation between a student's OLEffort and OLScore is non-negative vs. negative correlation.*

HB2: *The correlation between a student's OLScore and ICScore is non-negative vs. a negative correlation.*

Both are somewhat indirect ringer measurements. OLEffort corresponds to time and effort made which we assume a ringer would not undergo – a ringer would just do the assessments. A statistically significant negative correlation might indicate presence of the ringer effect (i.e., low OLEffort along with high OLScore) especially if a similar relationship holds between OLScore and ICScore (i.e., high OLScore but low ICScore).

HC: *A student's course grade can be predicted by the simple regression model*

$$CrsScore = \beta_0 + \beta_1 ICScore + \beta_2 OLEffort + \varepsilon$$

This was simply a check whether we could effectively bypass the ringer potential by eliminating the OLScore type of assignments. While there are significant issues to address, a strong predictive model might suggest we should not bother with online assignments and rely only on more easily validated measures.

STATISTICAL RESULTS

Details for Hybrid Courses (n=64)

HA: $H_0: \mu_{IC\ Score} - \mu_{OL\ Score} = 0$ vs $H_a: \mu_{IC\ Score} - \mu_{OL\ Score} < 0$

This hypothesis was tested with a two sample t-test (assuming unequal variances). The average ICScore is 77.5 and the average OLScore is 86.2; therefore, the difference, on average, is negative as opposed to the hypothesized no difference. The test statistic for the difference of IC-OL to test the null hypothesis is $t = -7.64$ with a t critical ($\alpha = 0.05$, one tail) of -1.66 . The statistical conclusion is to reject null and accept the alternate hypothesis and conclude that there is a difference in the average in-class score and on-line score.

HB1: $H_0: \rho_{(OLEffort, OLScore)} \geq 0$ vs $H_a: \rho_{(OLEffort, OLScore)} < 0$

The calculated correlation coefficient between these two variables is 0.1207. A t-test statistic was calculated to test the hypothesis. The computed test statistic is $t = 0.9417$ with a t critical ($\alpha = 0.05$, one tail) of -1.67 . The statistical conclusion is to fail to reject (FTR) the null which means we are not able to conclude the correlation coefficient is different than non-negative.

HB2: $H_0: \rho_{(OL\ Score, IC\ Score)} \geq 0$ vs $H_a: \rho_{(OL\ Score, IC\ Score)} < 0$

The calculated correlation coefficient between these two variables is 0.0260. A t-test statistic was calculated to test the hypothesis. The computed test statistic is $t = 0.2019$ with a t critical ($\alpha = 0.05$, one tail) of -1.67 . The statistical conclusion is to fail to reject (FTR) the null which means we are not able to conclude the correlation coefficient is different than non-negative.

HC: *The hypothesized multiple regression model postulates that Course Grade (\hat{Y}) is function of IC Score and OL Effort: $\hat{Y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, where x_1 is the OL Effort and x_2 is the IC Score.*

Ho: $\beta_1, \beta_2 = 0$ vs *Ha:* *At least one of the β 's differs from zero.*

The fitted model has an adjusted R^2 of 0.2068. Additionally, the F test statistic is 8.95 with a corresponding p-value of 0.0004. Based on this p-value we would reject the null and accept the alternate hypothesis and conclude at least one of the

coefficients is different from zero. Therefore, we would accept this model as a statistically valid model for predicting CRS score.

Online Courses (n=171)

HA: $H_0: \mu_{IC\ Score} - \mu_{OL\ Score} = 0$ vs *Ha:* $\mu_{IC\ Score} - \mu_{OL\ Score} < 0$

This hypothesis was tested with a paired t-test. The average ICScore is 86.6 and the average OLScore is 87.4; therefore, the difference, on average, is negative as opposed to the hypothesized no difference. The test statistic for the difference of IC-OL to test the null hypothesis is $t = -1.136$ with a t critical ($\alpha = 0.05$, one tail) of -1.66 . The statistical conclusion is to fail to reject (FTR) the null and conclude that we are not able to conclude the means are different.

HB1: $H_0: \rho_{(OL\ Effort, OL\ Score)} \geq 0$ vs *Ha:* $\rho_{(OL\ Effort, OL\ Score)} < 0$

The calculated correlation coefficient between these two variables is 0.1992. A t test statistic was calculated to test the hypothesis. The computed test statistic is $t = 2.642$ with a t critical ($\alpha = 0.05$, one tail) of -1.65 . The statistical conclusion is to fail to reject (FTR) the null which means we are not able to conclude the correlation coefficient is different than non-negative.

HB2: $H_0: \rho_{(OL\ Score, IC\ Score)} \geq 0$ vs *Ha:* $\rho_{(OL\ Score, IC\ Score)} < 0$

The calculated correlation coefficient between these two variables is 0.2818. A t-test statistic was calculated to test the hypothesis. The computed test statistic is $t = 3.818$ with a t critical ($\alpha = 0.05$, one tail) of -1.65 . The statistical conclusion is to fail to reject (FTR) the null which means we are not able to conclude the correlation coefficient is different than non-negative.

HC: *The hypothesized multiple regression model postulates that CRSScore (\hat{Y}) is function of IC Score and OL Effort: $\hat{Y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, where x_1 is the OL Effort and x_2 is the IC Score. *H₀: $\beta_1, \beta_2 = 0$ vs. *H_a: At least one of the β 's differs from zero.***

The fitted model has an adjusted R^2 of 0.8451. Additionally, the F test statistic is 167.48 with a corresponding p-value of $4.72 \text{ E-}25$. Based on this p-value we would reject the null and accept the alternate hypothesis and conclude at least one of the coefficients is different from zero. Therefore, we

would accept this model as a statistically valid model for predicting CRSScore.

INTERPRETATION

We obtained mixed results on HA for hybrid versus online courses. For hybrid courses, students did a statistically significant and material 8.9 points (on a 100 point scale) better online at a distance than on in class assignments which were simply randomized duplicates of the online assignments. In most cases, the students actually saw the online assignment first. For fully online courses the difference in performance online and in class was less than 1 point and not statistically significant.

Rejecting HA for hybrids supports the idea of ringers except neither of the correlations hypotheses provided confirming evidence. For hybrids, slightly positive but statistically insignificant correlations do not support the idea of ringers doing the distance work for students. For these courses, the components making up ICScore were not a significant portion of the course grade and we interpret our results as somewhat of a lack of interest or concern by students. Most of the grade in these courses came from development projects more interesting and relevant to the students.

For the fully online course, there was no evidence of ringers. In fact, the correlations were statistically significant positive suggesting that OLEffort, OLScore, and ICScore moved together as an instructor would hope. In these courses the ICScore was a significant part of the course grade and that 'hammer' seems to have forced learning to take place as designed in the curriculum.

For both types of courses, the regression (which tried to predict the course outcome but avoided any ringer influence) was statistically significant. For the online courses with its robust 85% R^2 , it is something of a rigged deal of course: most of the course grade came from the ICScore anyway. For hybrid courses, the 21% R^2 was less powerful but more telling. It raises the possibility that assessments where ringers cannot be ruled out may not be needed, though actually deploying this idea is another matter.

CONCLUDING REMARKS

Following a thorough study (but narrow rather than comprehensive) study, we have developed some reasonable assurance that our courses do not suffer significantly from ringer students doing online distance work for the actual students in the course.

Ringers are a chronic question for online courses or those with online components.

We are not surprised by these findings. We are fortunate to have a great deal of contact with students, even those doing the work online. Our classes tend not to be large or particularly anonymous, and there is significant instructor buy-in and active support. These have been established as key dimensions to success for online students.

For online courses, the significant, validated, in-class component probably played a role in ringer prevention: very little to be gained grade wise. For hybrid classes, there may have been a tendency to turn to ringers more often, but students may have seen the assignments as peripheral to the projects oriented work that makes up the bulk of these courses.

There are natural extensions including simply replicating the study. A structural equations approach including feedback on latent causal attitudes could be set up, but we feel this is not nearly worth the effort. More relevant and tractable would be research identifying the 'holy grail' of online instruction: what variables accurately measure online effort and commitment? Simply tracking time or mouse clicks put those who are lost and those who are exploring knowledgably and with alacrity in the same box. Pedagogical methods for bypassing the issue entirely along the lines of our hypotheses HC bears some investigation, but would not be appropriate in wide distance, totally anonymous instruction.

REFERENCES

1. Boehle, Sarah et. al. (2000). LearnNothing.com. *Training*, 37(11), 32.
2. Haythornwaite, C., Kazmer, M. M., Robbins, J., & Shoemaker, S. (2000). Community development among distance learners: Temporal and technological dimensions. *Journal of Computer Mediated Communications*, 6(1), 27.
3. Kanuka, K. (2002) Guiding principles for facilitating higher levels of web-based distance teaching and learning in post-secondary settings. *Distance Education*, 23(2), 163-182.
4. MacKinnon, G. R. (2000). The dilemma of evaluating electronic discussion groups. *Journal of Research on Computing Education*, 33(2), 125-131.
5. Özden, M.Y., Ertürk, I., & Sanli, R. (2004) Students' perceptions of online assessment: A case study. *Journal of Distance Education*, 19(2), 17-92.
6. Poole, D. M. (2000). Student participation in a discussion-oriented online course: A case study. *Journal of Research on Computing in Education*, 33(2), 162-177.
7. Ryan, M., Hodson-Carlton, K., Ali, S. (2005). A model for faculty teaching online: Confirmation of a dimensional matrix. *Journal of Nursing Education*, 44(8), 357-365.
8. Sautter, A. (2001). Evaluation of online students: methods and integrity issues. *University of Nebraska-Lincoln Dissertation*. Available online at <http://dwb.unl.edu/Diss/Sautter/>.